Topics in Modeling, Control, and Implementation in Automotive Systems

Gäfvert, Magnus

2003

Document Version:
Publisher's PDF, also known as Version of record

Topics in Modeling, Control, and Implementation in Automotive Systems

Topics in Modeling, Control, and Implementation in Automotive Systems

Magnus Gäfvert

Lund 2003

To my Mother (1943 – 1995) and Father and Dr. Tobben

Department of Automatic Control
Lund Institute of Technology
Box 118
SE-221 00 LUND
Sweden

ISSN 0280–5316
ISRN LUTFD2/TFRT--1066--SE

Printed in Sweden by Bloms i Lund Tryckeri AB.
Lund 2003

Abstract

This thesis treats different aspects of automotive control-systems and consists of four papers covering different areas in this large field.

The first paper presents a 9-degrees-of-freedom dynamic model of a tractor-semitrailer vehicle, aimed at simulation and evaluation of active chassis-systems for stability enhancements. Special focus is put on inclusion of the dynamics of load transfer, which are of importance in active yaw-control and roll-over prevention.

The second paper presents a new easy-to-use, semi-empirical tire model for combined-slip conditions. The model is based on combining the standard rigid-carcass brush model with available empirical pure-slip models. A new method is presented for partitioning empirical pure-slip forces into components of sliding and adhesive forces.

The third paper regards control of gasoline direct injection engines. Strategies based on feedback of the effective torque are shown to relax the requirements of accurate engine maps, in simulations. It is exemplified that extremum control may be used for online optimization of the engine operation with respect to fuel consumption. A new control structure that combines the extremum controller with disturbance rejection is used, to render the probing signals of the optimizing controller invisible in the engine output.

The fourth paper presents a new method to limit the effects of transient faults, occurring in hardware that hosts implementations of feedback controllers. The idea is to introduce artificial signal limits that are based on closed-loop signal bounds, and combine them with an anti-windup scheme.

Contents

Acknowledgments

I would like to thank all my colleagues at the department, who have provided with an outstanding environment of enthusiasm, creativity, expertise, generosity, and friendship, that made this work possible.

Particular thanks go to Karl Johan Åström, for encouraging me to start my Ph.D. studies and for advising me during my first years. Special thanks go to my advisor for the main part of the work, Björn Wittenmark, who has been a constant source of support and encouragement. Special thanks also go to Karl-Erik Årzén, for advising on parts of the work, and for being a good traveling companion. I am also in debt to Carlos Canudas-de-Wit (Laboratoire d'Automatique de Grenoble, ENSIEG-INPG, Grenoble) for giving me my first experiences of research work.

I thank all my co-authors, all with whom it has been a pleasure and honor to work. In specific, I would like to thank Lars Malcolm Pedersen, Olof Lindgärde (Volvo TD), Örjan Askerdal (Department of Computer Engineering, Chalmers), and Jacob Svendenius for close and stimulating cooperation. In the latter case combined with excessive candy-eating.

Many thanks go to Mats Andersson (Volvo TD) for guidance and fruitful discussions. All the weeks at Volvo TD were very inspiring and are much appreciated. I had very much fun and learned a lot working with Dr. Vilgot Claesson (Department of Computer Engineering, Chalmers) and Martin Sanfridson (Mechatronics Lab, Department of Machine Design, KTH). I will forever associate Göteborg with Thai food, Hong Kong action, and violent innebandy. Martin Törngren (Mechatronics Lab, Department of Machine Design, KTH) is much acknowledged for coordinating the DICOSMOS project and feeding it with many good ideas.

I would like to thank Serge Boverie and Dominique Passaquay (Siemens Automotive, Toulouse), for being helpful with the work on the engine control project.

Teaching and taking courses at the department have been just as enriching as doing research, and I wish to thank all the people who have contributed to this. Leif and Anders are much acknowledged for the outstanding computer support. Their skills are deep and mysterious. Likewise acknowledged is Rolf, for watching on us on the second floor and for programming the coffee-machine timer, and Eva, Britt-Marie, and Agneta for taking care of everything and everybody.

I wish to thank my former roommate Sven, for the excellent company and the innumerable discussions on things that were not related to research at all, at least not ours, and for putting up with me almost to the end. Thanks also go to Hubertus for helping me to fill my CD collection with strange jazz music. The Thursday innebandy gang has filled my legs with countless bruises. I owe you. I better not mention the Sundays.

I gratefully acknowledge the financial support for this work, received from The Swedish Agency for Innovation Systems (VINNOVA) under grants P1176-2 (DICOSMOS) and 2001-043 (LUCAS), Volvo Technological Development, and The Swedish Research Council (VR) under grant 2001-3020.

Finally, many warm thanks go to all my friends and family, for encouragement and much needed distraction. My last thanks go to you, Dr. Åsa, for your loving support, and for being there for me.

Magnus

1

Introduction

Control technology has increasingly become a vital part of modern passenger cars and commercial vehicles, starting with the introduction of electronic engine control and anti-lock braking systems in the late 1970s. It has been a necessary component in the development of low-emission engines with much improved fuel economy. It is the core technology in safety-enhancing functionality found in vehicles today, such as anti-lock braking, traction control, and active stability control. Other functionality has been realized to improve driving comfort and experience, such as air-conditioning, cruise-control, and active suspensions.

Many of these systems are based on old ideas, that could not be realized until the birth of cheap and reliable digital microcontrollers. The complexity of the algorithms involved, made analog implementations impractical and unreliable. Hence, the development of automotive control technology parallels the development of computer technology.

Even though new computer technology has enabled the realization of advanced control solutions in vehicles, the path to mature products is by no means a simple one. A road vehicle constitutes a process with many properties that make control difficult. Automotive applications, therefore, pose a number of great challenges to the control community:

- Control objectives that refer to driving experience may be difficult to formalize. For example, what is the desired response to a steering maneuver? Also, drivers are different in their use of the vehicle. Objectives may be conflicting, e.g., driving performance versus driving economy.
- The degrees-of-freedom are numerous, with six-degrees-of-freedom in the chassis movement for a single-body vehicle (longitudinal, lateral, heave, yaw, pitch, and roll motion), and additional degrees-of-freedom in the wheels, engine, axles, etc. The vehicle is a true multi-variable process, with complex couplings between actuator inputs and the controlled entities. For example, braking during a cor-

nering maneuver affects the yaw response of a vehicle. Most of the dynamics of the vehicle are also nonlinear.

- The vehicle is subject to significant parameter uncertainties due to load conditions, effects of aging and wear, road surface conditions, and temperature. Operating conditions may vary significantly and fast. Robust or adaptive techniques are imperative.
- Several functions are safety critical and must not fail under any circumstances. Functionality must be ensured with low maintenance. The environment is harsh and puts restrictions on sensor positions and hardware usage.
- The economic constraints are very strict, as the car manufacturers act on markets with huge competition.

These requirements would each by them self be a sufficient challenge and adding them together may seem overwhelming. Meeting them all would require a simultaneous control design for the entire vehicle. Traditionally, the approach has been different with a focus on control of separate subsystems individually. A problem with this is that undesired intercouplings between the different control schemes may occur, with degraded performance as a result. The trend for the future is therefore to take a more holistic perspective and integrate the control systems. This means a transfer from independent controllers to information-sharing cooperative designs. To date, this has been partly realized and made possible with the introduction of on-board vehicle computer-networks. A modern high-end car may have as many as 50 microcontrollers located on the vehicle body and connected by means of digital bus-systems.

The holistic approach to vehicle control is highly dependent on the availability of practical and accurate vehicle models for studying the effects of various control schemes. This includes models of the engine, transmission, steering system, suspension system, tires, chassis, sensors, actuators, and the vehicle data-buses. The traditional approach has been to model different subsystems individually. Therefore, special domain-specific modeling tools have been used for, e.g., electrical and mechanical systems, which makes integrated analysis of the full vehicle difficult. When systems interactions are the effects to study, these subsystem models need to be integrated.

Already today the safety requirements on the computer-controlled systems in vehicles are strict. The systems of tomorrow will pose even higher demands, with the introduction of new X-by-wire technology. There is a strive to remove mechanical linkages in the car and replace them with data buses and computer control. Among the reasons for this are increased performance, design flexibility, and reduced mass. The first example is

throttle-by-wire, which is now standard on many cars. Although unnoticed by most drivers, their accelerator pedal no longer has direct mechanical connection will the throttle plate. The first brake-by-wire passenger cars have just recently been launched on the market. They still have hydraulic brakes, but the valves are electronically controlled. In the near future electrical brake actuators will replace the hydraulic and pneumatic systems completely. The giant leap will be the introduction of steer-by-wire. Without this last mechanical linkage between the driver and the steered wheels, the directional response of the vehicle is entirely in the hands of computers. Such systems will therefore have extremely high requirements on dependability and fault tolerance.

In this context this thesis presents results in the areas of vehicle modeling, engine control, and fault-tolerant implementation of control algorithms. Topics that may seem diverse at first glance, but which are tightly related in the automotive industry.

1.1 Publications and Main Contributions

The thesis consists of four papers with introductory overviews. The papers and their main contributions are listed below.

Paper I

Gäfvert, M. and O. Lindgärde (2003): "A 9-DOF tractor-semitrailer dynamic handling model for advanced chassis control studies." *Vehicle System Dynamics*. Accepted for publication. Reprinted with permission from Swets & Zeitlinger Publishers.

This paper presents a flexible and modular 9-degrees-of-freedom nonlinear dynamic handling model for a tractor-semitrailer combination vehicle. The primary aim of the model is handling simulations with active yaw control, using unilateral braking. The model is of an intermediate detail level suitable for describing the dynamics of load transfer. The equations of motion are derived from the fundamental equations of dynamics in Euler's formulation, with the use of general computer-algebra software. The model may be used with standard simulation environments and a Simulink implementation is shown as an example. Simulation results are compared with experiments to validate the model. The article is based on [Gäfvert and Lindgärde, 2001, Gäfvert, 2001, Gäfvert *et al.*, 2000b].

Paper II

Gäfvert, M. and J. Svendenius (2003): "A novel semi-empirical tire-model for combined slips." Submitted to *Vehicle System Dynamics*.

The forces generated in the contact between the tires and the road are of major importance for the dynamic behavior of a road vehicle. Hence, accurate tire models are necessary components of complete vehicle models aimed at analyzing or simulating vehicle motion in real driving conditions. This paper presents a new method to combine empirical models for pure braking and cornering, to arrive at a model for simultaneous braking and cornering. The procedure is based on brush-model tire mechanics. The manuscript is based on the report [Gäfvert and Svendenius, 2003].

Paper III

Gäfvert, M., K.-E. Årzén, L. M. Pedersen, and B. Bernhardsson (2003): "Control of GDI engines using torque feedback exemplified by simulations." *Control Engineering Practice*. Accepted for publication. Reprinted with permission from Pergamon Press Ltd.

This article deals with a novel approach to the control of gasoline direct injection (GDI) engines. The control structure is hierarchical and relies on torque feedback as a central part. The sub-controllers are with a few exceptions designed using simple linear feedback and feedforward control-design methods, in contrast to traditional table-based engine control. A switching strategy, which maintains driving comfort during combustion-mode changes, is also proposed. A silent extremum-controller is used to minimize the fuel consumption in stratified combustion mode. The paper is based on the conference contributions [Gäfvert *et al.*, 2000, Gäfvert *et al.*, 2000a].

Paper IV

Gäfvert, M., B. Wittenmark and Ö. Askerdal (2003): "On the effect of transient data-errors in controller implementations." In *American Control Conference*, Denver, Colorado. Accepted for publication. Reprinted with permission from the American Automatic Control Council.

Computer-level faults leading to data errors in computations are predicted to occur increasingly frequent in future microprocessors. This paper discusses the impact of such errors on closed-loop performance in implementations of digital control systems. A method to render a control system more robust to data errors is presented and exemplified. Related work is presented in [Askerdal *et al.*, 2002, Claesson *et al.*, 2000, Sanfridsson *et al.*, 2000] and in the submitted journal paper [Askerdal *et al.*, 2003].

1.2 Other Publications

Several of the publications below provide with the basis for the manuscripts included in the thesis. The publications [Gäfvert, 1997, Panteley *et al.*, 1997, Panteley *et al.*, 1998, Olsson *et al.*, 1998, Gäfvert, 1999, Gäfvert *et al.*, 1999] treats dynamic friction modeling. They are related to Paper II, in that the specific model that was considered is now used for tire modeling by other authors.

Askerdal, Ö., M. Gäfvert, M. Hiller, and N. Suri (2002): "A control theory approach for analyzing the effects of data errors in safety-critical control systems." In *Proceedings of the Pacific Rim International Symposium on Dependable Computing*.

Askerdal, Ö., M. Gäfvert, M. Hiller, and N. Suri (2003): "Analyzing the impacts of data errors in safety-critical control systems." *IEICE Transactions on Information and Systems*, **E86-D:12**. Special issue on Dependable Computing. Invited paper. Submitted.

Claesson, V., M. Gäfvert, and M. Sanfridsson (2000): "Proposal for a distributed computer control system in heavy-duty trucks." Technical Report 00-16. Computer Engineering, Chalmers University of Technology, Sweden.

Gäfvert, M. (1997): "Comparisons of two dynamic friction models." In *Proc. Sixth IEEE Conference on Control Applications (CCA)*. Hartford, Connecticut.

Gäfvert, M. (1999): "Dynamic model based friction compensation on the Furuta pendulum." In *Proceedings 1999 IEEE Int. Conf. Control Applications and the Symp. Computer Aided Control Systems Design (CCA'99&CACSD'99)*. Kohala Coast, Hawaii.

Gäfvert, M. (2001): "Studies on yaw-control of heavy-duty trucks using unilateral braking." Technical Report ISRN LUTFD2/TFRT--7598--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gäfvert, M., K.-E. Årzén, B. Bernhardsson, and L. M. Pedersen (2002): "Control of gasoline direct injection engines using torque feedback: A simulation study." In Johansson and Rantzer, Eds., *Nonlinear and Hybrid Systems in Automotive Control*, chapter 13, pp. 289–320. Springer.

Gäfvert, M., K.-E. Årzén, and L. M. Pedersen (2000): "Simple linear feedback and extremum seeking control of GDI engines." In *Proceedings of Seoul 2000 FISITA World Automotive Congress*. Seoul, Korea.

Gäfvert, M. and O. Lindgärde (2001): "A 9-DOF tractor-semitrailer dynamic handling model for advanced chassis control studies." Technical Report ISRN LUTFD2/TFRT--7597--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gäfvert, M., L. M. Pedersen, K.-E. Årzén, and B. Bernhardsson (2000a): "Simple feedback control and mode switching strategies for GDI engines." In *SAE 2000 World Congress*. Detroit, Michigan, USA.

Gäfvert, M., M. Sanfridsson, and V. Claesson (2000b): "Truck model for yaw and roll dynamics control." Technical Report ISRN LUTFD2/TFRT--7588--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gäfvert, M. and J. Svendenius (2003): "Construction of novel semi-empirical tire models for combined braking and cornering." Technical Report ISRN LUTFD2/TFRT--7606--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gäfvert, M., J. Svensson, and K. J. Åström (1999): "Friction and friction compensation in the Furuta pendulum." In *Proc. 5th European Control Conference (ECC'99)*. Karlsruhe, Germany.

Olsson, H., K. J. Åström, C. C. de Wit, M. Gäfvert, and P. Lischinsky (1998): "Friction models and friction compensation." *European Journal of Control*.

Panteley, E., R. Ortega, and M. Gäfvert (1997): "An adaptive friction compensator for global tracking in robot manipulators." In *Proc. SYROCO '97, 5th IFAC Symposium on Robot Control*. Nantes, France.

Panteley, E., R. Ortega, and M. Gäfvert (1998): "An adaptive friction compensator for global tracking in robot manipulators." *Systems & Control Letters*, **33:5**, pp. 307–313.

Sanfridsson, M., V. Claesson, and M. Gäfvert (2000): "Investigation and requirements of a computer control system in a heavy-duty truck." Technical Report KTH/MMK/R--00/5--SE. Mechatronics Lab, Royal Institut, Stockholm, Sweden.

2

Vehicle Dynamics Modeling

A main objective of the control of vehicle dynamics is to improve the handling performance, or maneuverability, in order to obtain safer and more joyful driving. A range of systems are commercially available today, such as anti-lock braking, active stability-control, traction control, and 4-wheel steering, just to mention a few. Many new systems are under consideration, in parallel with efforts to integrate and coordinate the currently existing. In this process, models are necessary tools to analyze and design the control strategies and to evaluate the systems in simulations, possibly with hardware in the loop. An interesting approach in between simulation and full-scale vehicle testing is to use scale vehicles that preserve the dynamic characteristics [Brennan and Alleyne, 2001].

2.1 Model Requirements

Handling performance regard the planar motion of the vehicle and in many cases simple models that only include the planar degrees-of-freedom are sufficient. Analysis of simple bicycle models of the type shown in Figure 2.1 can explain much of the dynamic behavior in handling, particularly in steady-state [Wong, 2001]. In bicycle models, the wheels on each axle are usually lumped together to an equivalent single-track configuration. The chassis motion is described by the longitudinal and lateral velocities, and the yaw rate. The vehicle side-slip is often used instead of the lateral velocity. Linearized equations of motion at constant speed are very simple and describe the vehicle behavior fairly well in normal driving conditions. The tire forces, which determine the chassis motion, may then be described linearly in terms of the vehicle states. For driving in limit conditions the tires need to be described nonlinearly.

When the chassis is subject to acceleration, in cornering, braking, or driving, the vehicle weight is transfered between the wheels. The dynam-

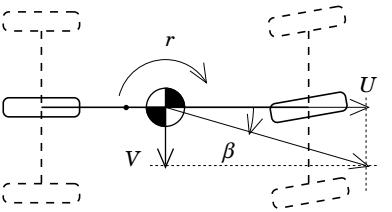

Figure 2.1 The single (double) track bicycle model. Vehicle motion is described by the longitudinal, U, and lateral, V, velocities, and the yaw rate, r. The vehicle side-slip β is often used instead of the lateral velocity.

ics of this load transfer are in large determined by the suspension characteristics. The load transfer has effects on the tire forces and must therefore be included at large accelerations. A quasi-static approach is to use the effective axle-characteristics [Pacejka, 2002]. If transient behavior is of interest, then the roll motion and, possibly, also pitch and bounce, must be described dynamically. For a single-unit vehicle this means that up to six degrees-of-freedom must be included, and the equations of motion become somewhat more involved. Examples of models are found in [Pacejka, 2002, Kiencke and Nielsen, 2000]. Models with more than three degrees-of-freedom for the chassis motion are mostly used for simulations, as they are complicated to analyze.

In addition to the chassis equations-of-motion, a vehicle model needs to describe various subsystems. The most important is the tires, which have critical influence for all chassis-control systems. Linear tire models, when applicable, are normally lumped with the equations of motion. Wheel dynamics may be included but models that describe tire forces apart from wheel rotation are also possible. When out-of-plane freedoms are included for the chassis also the suspensions must be described. These introduce new nonlinearities in the model. Depending on the application also models of the braking system, steering system, the powertrain, etc., need to be included.

Simulations for evaluation of chassis-control systems are often performed with standardized test maneuvers, such as J-turns and severe lane-changes. This may either regard open-loop responses, or with a driver model included. Advanced realtime simulators may take input from a human driver, or allow for the introduction of hardware in the loop. This is a very useful method to evaluate third-party subsystems such as anti-lock braking systems, etc. Realtime simulation requires efficient models with respect to the use of computational resources.

2.2 The Tractor-Semitrailer Vehicle

Tractor-semitrailer vehicles of the type illustrated in Figure 2.2 are very common in commercial fleets. Chassis control for these vehicles mainly

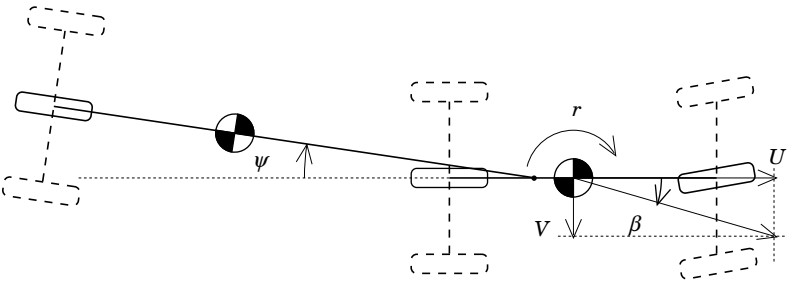

Figure 2.2 The single (double) track bicycle model for a tractor-semitrailer vehicle. Vehicle motion is described by the longitudinal, U, and lateral, V, velocities, the yaw rate, r, the articulation angle, ψ, and the articulation rate, $\dot{\psi}$. The vehicle side-slip β may be used instead of the lateral velocity.

regard safety. Due to the high center-of-gravity with respect to track width, rollover may occur at lateral accelerations that exceed certain limits. Rollover has been identified as one main cause of accidents for heavy-duty vehicles [Sampson, 2000]. Another common cause of accidents for this dual-unit vehicle is so called jackknifing, which may occur very rapidly if the rear wheels of the tractor loose grip [Palkovics and Fries, 2001]. Much research has been devoted to control systems that prevent these hazardous situations. Recent investigations, e.g. [Hac, 2002], show that the means for yaw stabilization, related to jackknifing, and rollover prevention are partly coinciding. One reason for this is that vehicle oversteering is a contributing factor to rollover.

The high center-of-gravity implies that roll, with corresponding load transfer, is significant already at moderate accelerations. This, in combination with the fact that rollover limits are difficult to describe statically, motivates the use of models with at least roll freedoms [Dahlberg, 2000]. In addition, this dual-unit vehicle introduces the semitrailer articulation as a new freedom in the plane. Due to the kinematic constraints of the hitch, in total nine degrees-of-freedom are needed to describe translation, yaw , pitch, and roll motion of the chassis.

2.3 Active Stability Control

The tire-road friction limits restrict the envelope of stable driving conditions. If friction limits are reached for the tires on an axle, the tire forces saturate. The dynamic behavior of the vehicle then changes, which possibly leads to hazardous situations. Saturation on the front axle results in understeering and on the rear axle of a single-unit vehicle in oversteering, with possible spin-out. On a tractor-semitrailer, rear-axle saturation may lead to jackknifing and semitrailer-axle saturation to so called trailer swing.

The basic principle of active stability control (ASC) systems is to calculate a desired vehicle yaw-rate, using the vehicle speed and the steering-wheel angle commanded by the driver. Using feedback from a yaw-rate sensor, a compensating moment is computed, which is then mapped into unilateral braking commands. Oversteering is normally compensated by braking the outer front wheel and understeering by braking the inner rear. These braking actions produce corrective moments on the vehicle from the longitudinal braking forces, but also from resulting reductions of the lateral tire-forces. On tractor-semitrailer vehicles also the trailer brakes are applied at oversteering, to straighten the vehicle combination so that jackknifing is avoided. If only the yaw-rate is controlled then the vehicle may describe the desired yaw-motion without following the corresponding curvature path. To ensure the path-following it is necessary to keep the orientation of the vehicle aligned with the direction of the vehicle velocity, which is equivalent to keeping the vehicle side-slip small. An additional control objective for the ASC system is, thus, to limit the side slip and most systems are based on feedback also from this variable.

In ASC for tractor-semitrailer vehicles also the articulation angle and rate may be subject to control. In current production systems, however, the tractor is controlled more or less as a single unit vehicle [Hecker *et al.*, 1997].

It is already noted that reduction of oversteering also reduces the risk of rollover. Even better prevention of rollover is obtained by combining the ASC with rollover detection [Palkovics *et al.*, 1999]. Then, when rollover is imminent, brake actuation is used to reduce the vehicle speed and yaw-rate and, hence, the lateral acceleration. The reduction in lateral tire-forces resulting from the braking will also help.

2.4 Paper I

Bicycle models of the type shown in Figure 2.2 may be used to analyze and simulate yaw stability on low-friction surfaces. Then, the lateral ac-

celeration is limited beacause of the small lateral tire forces and the load transfer may be neglected. In [Gäfvert, 2001] such models were used to investigate different strategies for yaw-stabilization based on LQ feedback from tractor and semitrailer states. The main advantage compared to the commonly used feedback from only tractor states is that also the risk of trailer swing appears to be reduced.

The 9-degrees-of-freedom tractor-semitrailer model presented in Paper I is aimed at simulations to evaluate active stability-control systems of the type discussed, in scenarios where load transfer has significant effect.

3

Tire Models for Handling Simulations

On all road vehicles the chassis-movement is actively controlled by steering, throttling, braking, and possibly suspension actuations, which all ultimately result in forces in the tire-road contact. The tires are, therefore, components with major influence on the dynamic behavior of a vehicle. Obviously, accurate tire models are necessary parts of models aimed at analyzing or simulating vehicle motion in real driving conditions. With new active chassis-control systems that are based on unilateral braking it is increasingly important to describe the effects of combined braking and cornering correctly, also near and beyond the tire-road friction limits.

3.1 Understanding the Tire

The now predominating radial tire is built on a carcass of steel cords running radially from bead to bead [Bauer, 1996], see Figure 3.1. The beads clamp the tire to the wheel rim. A stabilizing belt of crossed steel-cords surrounds the carcass. The rubber tread is bonded on the belt and sculpted with a tread pattern. In the contact patch between the tire and the road, the rubber is partly gripping and partly sliding on the road surface, resulting in forces of both static and kinetic friction. Hence, different physical processes are present simultaneously in different regions of the contact patch. The generated forces under different driving conditions depend on the characteristics of the friction interface between the tire and the road, the motion of the tire carcass relative to the road, and the applied normal load. There are various ways to describe this motion, depending on choice of reference systems and normalizations. The most common is to follow the standards of the Society of Automotive Engineers (SAE) [SAE Recom-

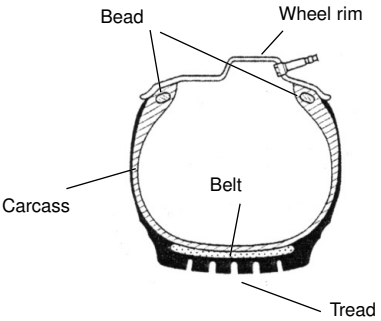

Figure 3.1 Schematic of a radial truck tire. (Reprinted from [Bauer, 1996].)

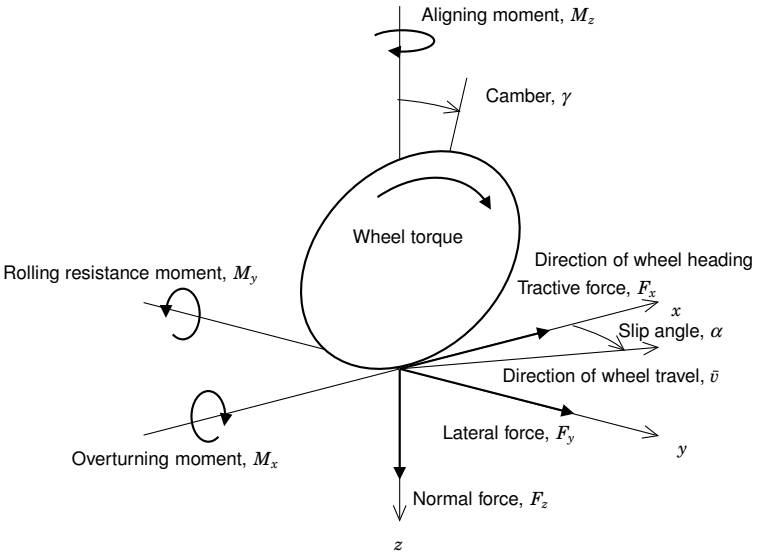

Figure 3.2 Forces and moments acting on a tire [SAE Recommended Practice J670e, 1976].

mended Practice J670e, 1976], with the longitudinal x-axis aligned with the wheel heading, the lateral y-axis perpendicular to the wheel, and the vertical z-axis pointing downwards, as in Figure 3.2. In vehicle-handling studies, the lateral and longitudinal planar forces, F_x and F_y, and the self-aligning moment, M_z, are of interest. The longitudinal tire force is

generated when braking or driving, and the lateral force and the torque when cornering. The self-aligning moment results from the fact that the planar forces have a point of action that is not positioned exactly under the wheel center. The rolling resistance and overturning moment are not of primary interest for vehicle handling.

Tire Kinematics

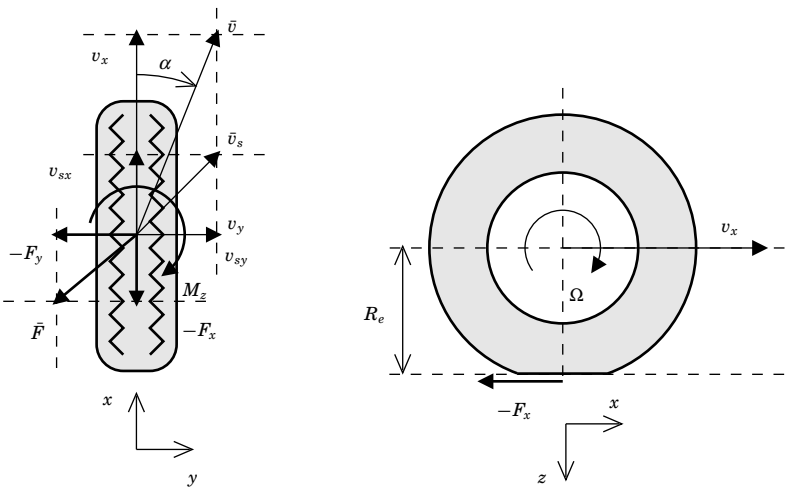

Figure 3.3 Kinematics of a tire during braking and cornering. Force vectors are also shown. (Left: top view; Right: side view)

The tire motion is determined by the *wheel-travel velocity*, $\bar{v} = (v_x, v_y)$, and the *circumferential velocity* of the tire, $v_c = \Omega R_e$, where Ω is the wheel angular velocity and R_e the effective rolling radius. These variables are not very practical for describing tire forces and moments. The *slip velocity*, $\bar{v}_s = (v_x - v_c, v_y)$, is the relative motion of the tire to ground in the contact patch. In tire modeling, the slip velocity is often normalized with a reference velocity, to yield a dimensionless *tire-slip* vector. It is the custom to describe tire forces as functions of the slip, rather than on the slip velocity. Implicitly, this assumes that the forces do not depend on the magnitude of the slip velocity, v_s. In general, this is not true since at least the contributions from sliding friction normally depend on the velocity. The most commonly used slip is defined with the longitudinal component of the wheel-travel velocity as the reference velocity. The longitudinal component of the tire slip is called the *slip ratio*

$$\lambda = \frac{v_x - v_c}{v_x} \tag{3.1}$$

which is often expressed in percent. The lateral component is

$$\tan(\alpha) = \frac{v_y}{v_x} \tag{3.2}$$

where the *slip angle* α, normally in degrees, is the deviation of the wheel-travel velocity from the wheel heading. The slip ratio and the slip angle are the standard variables that are used to describe the tire motion.

Tire Mechanics in the Small Scale

To understand the qualitative behavior of pneumatic tires it is good to have a picture of the physical processes that occur during driving. First, regard that there is a contact region of a certain size between the tire and the road. The normal load on the tire is carried by a near Hertzian pressure distribution longitudinally along this region. Rubber tread-elements on the tire circumference pass through the contact region as the tire rolls. At free rolling, the elements adhere to the road surface from entering the region, until leaving the rear end.

If there is a longitudinal slip, or equivalently, a longitudinal slip-velocity, then an adhering element is constrained to being fixed to ground, while the tire has a relative motion. This will induce a local deformation, which grows larger with the time elapsed since an element entered the contact region. Since the elements are traveling through the region by the wheel circumferential velocity, this means that the deformation grows larger with the distance from the leading edge. The local deformation gives rise to local (visco-)elastic forces, which depend on the tread stiffness, and in total a force distribution along the contact region. At a certain distance from the leading edge the local force exceeds the static friction limit and the tread elements break away and start to slide. From this point to the rear end of the contact region the local forces are described by kinetic friction and depend on the pressure distribution and the kinetic-friction coefficient, which is, in general, velocity dependent. In summary, the total tire force is composed of the deformation force from the adhering region, stretching from the leading edge to the break-away point, and of the kinetic friction force from the sliding region, stretching from the break-away point to the trailing edge. The break-away point moves from the trailing edge, at free rolling, to the leading edge with increasing slip magnitude. At a certain slip magnitude the point reaches the leading edge and full sliding occurs.

In the lateral dimension the process is analog. If there is a slip angle, or equivalently, a lateral slip velocity, an adhering element will deform laterally as it travels through the contact patch. If all tire deformation was located in the tread, then the tread-element deformation would be

described by $\tan(\alpha)$ times the distance from the leading edge, with a corresponding deformation force distribution. However, a difference to the longitudinal case is that deformation of the carcass will act such that only part of the slip angle results in tread deformation. The asymmetrical distribution of the lateral force with respect to the wheel-center projection point on ground, results in the self-aligning moment that strives to turn the wheel around the vertical axis.

If longitudinal and lateral slips are present simultaneously, then the longitudinal and lateral processes are superimposed. Both the longitudinal and lateral deformation forces are then carried by the static friction. As a result, the break-away point is closer to the leading edge than if each slip component were present individually. The force from the sliding region is counter-directed to the slip velocity. The orientation of the force from the adhering region deviates from the sliding-force orientation, depending on the relative longitudinal and lateral tread stiffnesses and on the carcass deformation. Therefore, as the adhering region shrinks with increasing slip magnitude, the orientation of the force vector will change until it is opposed to the slip-velocity at full sliding. Carcass deformation at combined slip also has influence on the self-aligning moment.

The principles sketched above give a rough picture of the mechanisms in the tire-road interface. In addition to this, there are several important factors that influence the tire behavior. Running on dry or wet asphalt, gravel, snow, or ice give different behavior. Tread wear changes the characteristics. Tread and road-surface temperatures also have substantial influence. The dynamics of carcass deformation have effetcs on the transient behavior.

In vehicle simulation, only the macro-level behavior of tires is of importance. However, understanding of the small-scale mechanics makes it easier to understand the behavior on a larger scale and also to understand the principles behind various tire models.

Behavior at Pure Braking or Cornering

At pure braking, the slip angle is zero and the resulting longitudinal tire-force, F_{0x}, is determined solely by the slip ratio. In Figure 3.4 (left), the longitudinal tire force at different slip ratios in pure braking is shown for a typical truck tire on a steel surface. As can be seen, the overall characteristics are highly non-linear. For small slips the longitudinal force appears to be approximately linear. Then, a peak appears around $\lambda = 20\%$. After the peak, the force level out and decreases with increasing slips. The shape of the curve is due to the adhesive region of the contact patch getting smaller with increasing slips. At a certain slip there is no adhesion at all and the force is generated entirely by sliding friction. The transition to full sliding occurs in the neighborhood of the peak value.

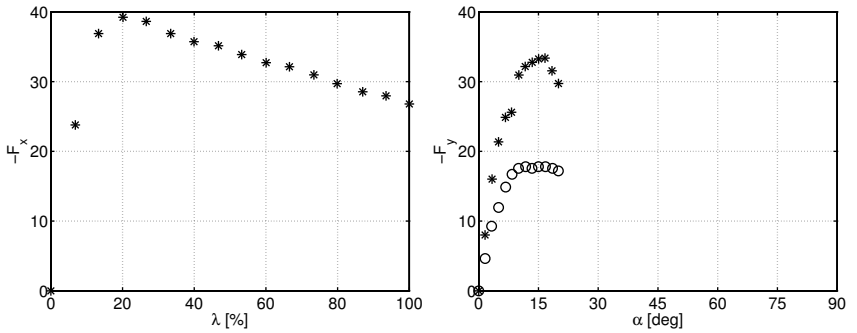

Figure 3.4 Measured tire forces of a standard truck tire on steel surface at F_z = 20 ("o"), and 40 ("*") kN.† Left: Longitudinal force at pure braking/traction. Right: Lateral force at pure cornering. (Data from [Edlund, 1991].)

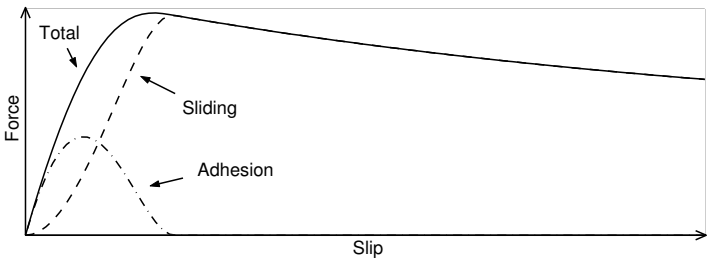

Figure 3.5 As the tire is partly gripping and partly sliding on the road surface the total tire force is composed of adhesive and sliding forces.

In pure cornering the slip ratio is zero, since the tire is rolling freely, and the resulting lateral tire-force, F_{0y}, depends only on the slip angle. Figure 3.4 (right) shows the lateral tire force for different slip angles at pure cornering. The shape is qualitatively the same as for the longitudinal force and reflects the transition from adhesion to sliding in the contact patch. The force peak is found around $\alpha = 15$ deg.

In Figure 3.5 is it illustrated how the total tire force has components resulting from the adhesive and sliding parts of the contact region. The behavior described so far concerns steady-state conditions. For varying slips there are transient effects that results from carcass deformation. In princip, the tire act as a complex rolling spring when attached to a vehicle. The main observed effect of the transient behavior is that it takes a specific rolling distance, the *relaxation length*, σ_c, for the lateral slip, and hence, the lateral force, to build up.

†The longitudinal force, F_x, was only measured for one normal load.

Behavior at Combined Braking and Cornering

At combined braking and cornering the magnitude and direction of the resulting tire force depends on the magnitude and direction of the tire-slip vector. Figure 3.6 shows the resulting tire forces for two sets of total slips where the lateral slip is held fixed and the longitudinal slip is varied from free rolling ($\lambda = 0\%$) to locked wheel ($\lambda = 100\%$). The resulting lateral force decreases, compared to the pure cornering case, as the applied longitudinal slip increases. This effect may be explained by the decreasing size of the adhesive contact region, as the total slip magnitude increases, and by the projection of the friction force that is generated in the sliding region on the longitudinal and lateral directions. The maximum available tire forces, F_{0x}^* and F_{0y}^*, at pure braking or cornering, reflects the maximum available tire-road friction, which may be, in fact, a mix of static and kinetic friction. Since the friction provides a bound also on combined forces, the envelope of all possible curves as in the figure describes an ellipse-like shape, with corresponding major and minor axes defined by F_{0x}^* and F_{0y}^*. This envelope is commonly referred to as the *friction ellipse*.

Dependence on Normal Load

The tire-forces depend nonlinearly on the normal force, F_z. This dependence is due to, e.g., changes in the in the area of the contact region, changes in the pressure distribution, and changes in material properties. Effects of the normal-load on the lateral force are illustrated in Figure 3.7. In general, the initial slope and the peak value of the normalized force, F_{0y}/F_z, decrease with increased load.

Tire Measurements

Measurements of tire characteristics are performed with test facilities where parameters such as the tire slips, normal load, and wheel travel velocity are controlled precisely while forces and moments are recorded. A common setup is to apply the loaded tire on a large, rolling drum [Holdmann *et al.*, 1999]. Another method is to translate a flat steel bar under the loaded tire [Nordström, 1993]. Testing under real road-surface conditions may be performed with a test rig attached to a vehicle [Pottinger *et al.*, 1993]. Some test facilities are limited to pure-slip or steady-state conditions, while others may be used to measure combined-slip conditions and transient behavior. Tests are normally performed by holding one slip constant, while sweeping the other. In general, the applied side-slip angles are smaller than around 20 degrees and empirical tire-force data is normally only available up to this limit. Resulting data are of the type shown in Figures 3.4 and 3.6.

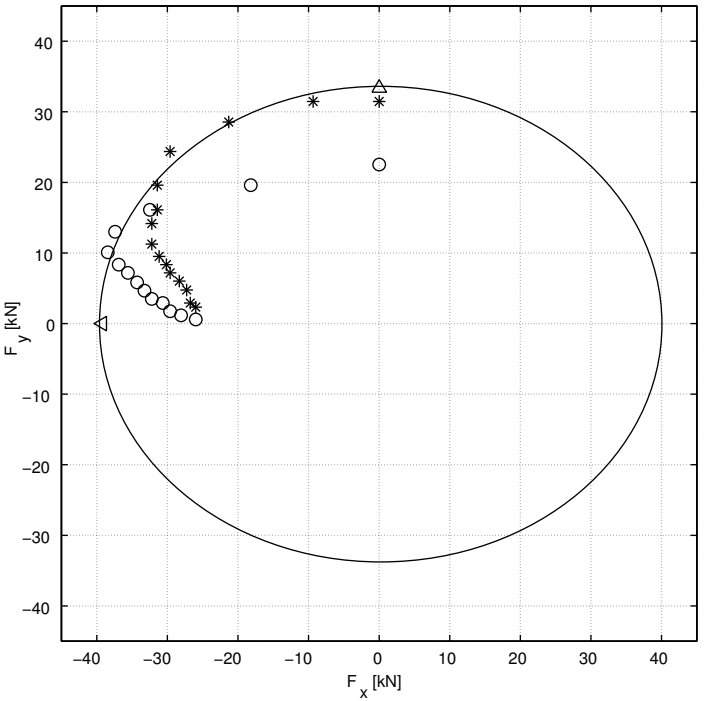

Figure 3.6 Measured combined tire forces of a standard truck tire on steel surface at $F_z = 40$ kN and $\lambda \in [0, 1]$ for $\alpha = -4.7°$ ("o"), and $\alpha = -9.8°$ ("∗"). The triangles denote the maximum achievable pure longitudinal and lateral forces (F_x^*, F_y^*) and defines the friction ellipse (dashed). (Data from [Edlund, 1991].)

3.2 Practical Tire Models

The basic demands on practical tire models for simulation of vehicle handling are the abilities to reproduce forces and moments from compact and comprehensible mathematical expressions, using parameters that give a compact representation and are easily obtainable. Since the physical processes occurring in the contact region are difficult to model theoretically, the resulting forces and moments are often described in terms of empirical data, tire slips, and normal forces. The empirical data inherently include the complex effects of friction, roughness, lubrication, deformation, etc.

Pure-Slip Models

Since the characteristics of the tire are nearly linear for small slips the

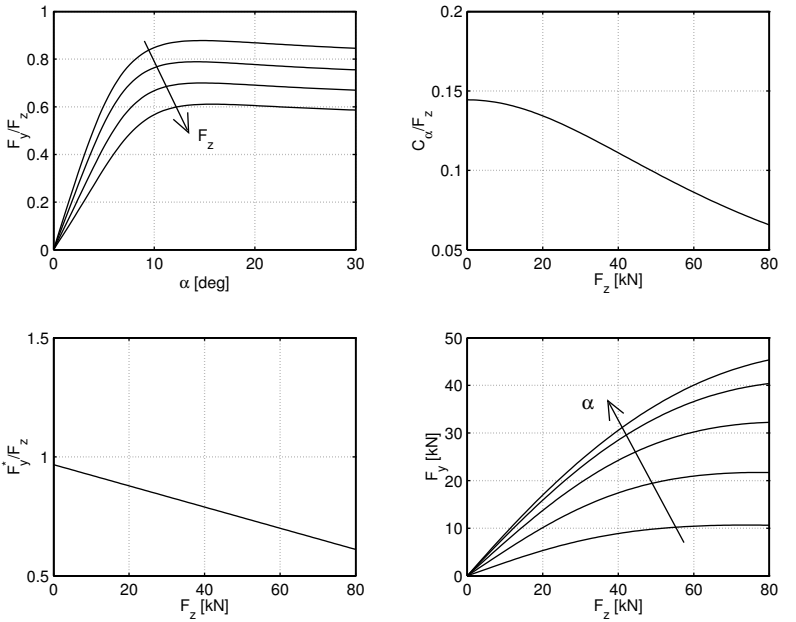

Figure 3.7 Dependence on normal load for a typical truck tire. Upper left: normalized pure-slip lateral force; upper right: normalized cornering stiffness; lower left: normalized peak-value; lower right: variation of pure-slip lateral force with normal load for fixed side-slips. (Based on data from [Edlund, 1991].)

forces at pure braking or pure cornering may be described by

$$F_{0x}(\lambda) = -C_\lambda \lambda \tag{3.3a}$$

$$F_{0y}(\alpha) = -C_\alpha \alpha \tag{3.3b}$$

where the braking stiffness, C_λ, and cornering stiffness, C_α, are defined as the linearization of the friction curves in Figure 3.4 at $\lambda = 0$ and $\alpha = 0$ [Wong, 2001]

$$C_\lambda = -\left.\frac{dF_{0x}}{d\lambda}\right|_{\lambda=0} \tag{3.4a}$$

$$C_\alpha = -\left.\frac{dF_{0y}}{d\alpha}\right|_{\alpha=0} \tag{3.4b}$$

A common way to include the transient effects resulting from the flexible carcass is to replace Equation (3.3b) with the linear time-variant differ-

ential equation

$$\frac{\sigma_c}{v_c}\frac{dF_{0y}}{dt} + F_{0y} = C_\alpha \alpha \tag{3.5}$$

where the time constant depend on the circumferential velocity, v_c, and the relaxation length, σ_c, [Pacejka, 1988].

For larger slips the nonlinear characteristics of the friction curves need to be modeled. The now predominating model is the "Magic Formula" [Bakker *et al.*, 1987], which is a functional approximation of the lateral and longitudinal tire forces, as well as aligning torque, on the form

$$f(\xi) = D\sin\Big(C\arctan\big((1-E)\xi + (E/B)\arctan(B\xi)\big)\Big) \tag{3.6}$$

where (ξ, f) is (λ, F_{0x}), (α, F_{0y}), or (α, M_{0z}). The four coefficients have interpretations as stiffness factor (B), shape factor (C), peak factor (D), and curvature factor (E), and are unique for each of F_{0x}, F_{0y}, and M_{0z}. Approximate normal load dependence may be introduced as $D = a_1F_z^2 + a_2F_z$, $BCD = (a_3F_z^2 + a_4F_z)/e^{a_5F_z}$, and $E = a_6F_z^2 + a_7F_z + a_8$, or with the slightly modified expressions of [Bakker *et al.*, 1989]. Fit to experimental data is performed by parameter optimization of B,C,D,E, or $a_{1..8}$. Another related model is the Burckhardt approach [Kiencke and Nielsen, 2000], in which another functional approximation is used.

Combined-slip models

Models for combined-slip forces may roughly be divided into those that are based on manipulation of empirical pure-slip forces, those that offer a compact mathematical representation of combined-slip empirical data, and those that are based on physical models using empirical parameters.

Models Based on Manipulation of Pure-Slip Forces. A very simple model of combined forces is based on the friction ellipse illustrated in Figure 3.6 [Wong, 2001]. For a given longitudinal force, F_x, an elliptic constraint is solved for the combined slip lateral force, $F_y(\lambda, \alpha)$. The constraint is defined by major and minor axes given by the peak value of the longitudinal pure-slip force, F_{0x}^*, and the lateral pure-slip force, $F_{0y}(\alpha)$. The result is that the lateral force is described as a function of the longitudinal force, which is not correct for large slips, as is evident from the figure. A drawback is also that the longitudinal force is assumed to be known. Another simple model is the Kamm Circle [Kiencke and Nielsen, 2000], where the resultant force magnitude is described as a function of the total slip magnitude. The force and slip vectors are then assumed to be collinear, possibly with a corrective factor. A drawback with this model is that longitudinal and lateral characteristics are assumed to be

the same, modulo the corrective factor, and that the orientation of the force is not slip dependent. Some other early efforts to model tire forces under combined-slip conditions are described and compared in [Nguyen and Case, 1975]. One of the most well-known is presented in [Nicholas and Comstock, 1972]. Here the combined slip forces are described by the projection of the pure slip forces on the vector $(F_{0y}(\alpha)\lambda, F_{0x}(\lambda)\tan(\alpha))$. In [Brach and Brach, 2000] this model is shown to give incorrect result for small slips and a modified version is presented. In the recent COMBINATOR model [Schuring *et al.*, 1996], the combined-slip slip-magnitude is used with empirical pure-slip models. The force magnitude is then given by a simple ellipse-like functional approximation, based on the pure-slip forces. The resulting force is assumed to be collinear with the slip vector at all slip magnitudes. In [Bakker *et al.*, 1987] a procedure for computing combined forces based on Magic Formula pure-slip models is presented. The use of a normalized slip ensures that the pure-slip models are used at points corresponding to the same size of the adhesion region, which is sensible from a physical point of view. A slip dependent correction factor is used to obtain the correct orientation of the force vector over the entire slip range. It is not clear how to determine this correction factor and in [Bakker *et al.*, 1989] a modified procedure was presented, that essentially introduces a compact and physically sound parameterization of the force orientation.

Models Based on Empirical Data Reduction. In [Bayle *et al.*, 1993] a model is presented that is based on functional representation much inspired by the Magic Formula. Functions based on arc tangents are used to describe forces under combined-slip conditions. In principle, these represent purely empirical weighting functions that are multiplied with pure-slip forces obtained with the Magic Formula. The latest formulation may be found in [Pacejka, 2002] and is also commercially available in the "Delft-Tyre" product series of TNO Automotive, Netherlands. The model needs to be calibrated with combined-slip data.

Physical Models. Models based on physical first principles have the advantage that physical parameters provide a compact representation of tire and road characteristics. However, models of considerable complexity must be applied to catch all phenomena of interest and all required parameters are not easily obtainable. Such models are, therefore, not suited for simulation of vehicle dynamics where short simulation times are often desired.

Less complex models that catch the main effects may also be formulated. The widely used brush model describes the generation of tire forces by regarding the tread as consisting of infinitesimal bristles. A commonly

used model is described in [Segel, 1984], where an assumption of rectangular normal-pressure distribution is combined with the bristle mechanics. The model is based on parameters that describe tire stiffnesses, normal load, slips, and road friction coefficients. The assumption on the pressure distribution leads to simple expressions, but qualitatively less correct results. In [Gim and Nikravesh, 1991] a parabolic pressure-distribution is used instead. Still, the simplifying assumptions restrict the practical use of the model, since effects of, e.g., velocity dependence and carcass deformation are neglected.

Other Models. Another class of models that have gained much attention recently are those that include also transient behavior. The Magic Formula, and other similar models, are approximations of experimental steady-state tire characteristics. The main dynamic effect is due to the flexible carcass, which in the simplest form may be modeled by (3.5). Dynamic transients were investigated in, e.g., [van Zanten *et al.*, 1989] and some recent tire models aim at including dynamic behavior [Bliman *et al.*, 1995, Canudas De Wit and Tsiotras, 1998, Deur, 2002]. The state-of-the art tire model that includes transient characteristics is the SWIFT model based on [Maurice, 2000] and [Zegelaar, 1998], also commercially available as a "DelftTyre" product.

A fair amount of research has been devoted to tire modeling and this brief survey by no means cover the area. A version with more technical information is given in [Gäfvert and Svendenius, 2003]. More complete information may be found in the outstanding reference on tire modeling [Pacejka, 2002].

3.3 Paper II

State-of-the-art tire models provide with more than enough accuracy for most vehicle dynamics simulations. In practical use they require the use of commercial packages due to their complexity and their dependence on accurate and numerous parameters. For users with intermediate requirements it is of interest to have easy-to-use models that may be implemented by the end user, with reasonable effort. An additional desire is to keep the number of parameters and the required base of empirical data as small as possible. Based on the mechanics of the tire and observations from empirical data, a number of criteria for such models may be stated (in the spirit of [Brach and Brach, 2000]):

- The combined force $\bar{F}(\lambda, \alpha)$ should preferably be constructed from pure slip models $F_{0x}(\lambda)$ and $F_{0y}(\alpha)$, with few additional parameters.

This gives a compact description of tire characteristics from easily obtainable data.

- The computations involved in the model must be numerically feasible and efficient.
- The formulas should preferably be physically motivated and the parameters have clear physical interpretations.
- The combined force $\bar{F}(\lambda, \alpha)$ should reduce to $F_{0x}(\lambda)$ and $F_{0y}(\alpha)$ at pure braking or cornering:

$$\bar{F}(\lambda, 0) = (F_{0x}(\lambda), 0)$$
$$\bar{F}(0, \alpha) = (0, F_{0y}(\alpha))$$

- Sliding must occur simultaneously in longitudinal and lateral directions.
- The resulting force magnitudes should stay within the friction ellipse, since this constitutes a physical constraint.
- The combined force should be $\bar{F} = -(C_\lambda \lambda, C_\alpha \alpha)$ for small slips and $\bar{F} = -F_z \mu \bar{v}_s / v_s$ at full sliding for tires with isotropic friction characteristics.

Guided by these criteria a new model has been derived, which is presented in Paper II. The new model belongs to the class of models that construct combined-slip forces from pure-slip data. However, while many of the earlier models are based on purely empirical formulas, the proposed model is based on physical relations that are derived from brush-model mechanics. The result is a model with the physical foundation of the brush model, that still incorporates effects of load dependence, velocity variations, and to some extent also steady-state carcass deformation. The proposed model is one in a family of models derived in [Gäfvert and Svendenius, 2003].

4

SI Engine Control

Electronic control of spark ignition (SI) engines made the cars of the 1990s ten times cleaner and twice as fuel efficient as the cars of the 1970s [Powers and Nicastri, 2000]. Emission requirements keep evolving and new techniques are searched to comply with ever stricter regulations. Control technology is a key factor in this development but also in the integration of the engine into vehicle-level control-systems. In modern vehicles the engine is an actuator that may receive commands not only from the driver, but also from other systems.

In the first years of electronic engine control the design concept was focused on controlling the auxiliary devices such as fuel supply and ignition systems, rather than the full engine. Today all engine-related control functions, denoted engine management, are normally integrated into a single electronic control unit. Demands on emission and fuel-economy require control of several engine variables such air/fuel ratio, ignition timing, exhaust-gas recirculation (EGR), for all engine speeds, loads and temperature conditions. Engine management systems (EMS) are, in principle, partitioned into control of fuel injection and control of ignition. The first deals with forming a suitable mixture of air and fuel to charge the cylinders and the other with igniting this mixture at a suitable crank-angle position. The operation of standard engine-control systems is surveyed in [Chowanietz, 1995]. More technical descriptions of control schemes are described in [Kiencke and Nielsen, 2000].

4.1 Lambda Control

The motivating factors for the development of electronic engine controls have been the increasingly strict governmental requirements on emission of carbon monoxide (CO), oxides of nitrogen (NOx), and unburned hydrocarbons (HC). The first step in the late 1970s was to introduce 2-way

catalysts that could remove CO and HC. EGR was then applied to reduce NOx levels. As requirements on NOx limits went stricter, the 3-way catalyst was introduced. This catalyst can operate with only a very small amount of residual oxygen present in the exhaust gas. Therefore, the ratio of air to fuel used in the cylinders need to be controlled to stoichiometry, that is such that precisely the amount of oxygen that is needed to burn the fuel is injected. This was solved by the exhaust-gas oxygen-sensor (EGO), or lambda sensor, in combination with electronic control. The EGO sensor effectively has a relay characteristic in detecting excess oxygen. Then, in principle, a PI-controller uses the sensor signal to direct the injected amount of fuel, to achieve stoichiometric combustion. Since the sensor is located downstream the fuel injection there is a transport delay present in the control-loop. This, in combination with the nonlinear sensor characteristics, results in a limit cycle and the air-fuel ratio normally oscillates with a frequency of 0.2–2 Hz under normal driving.

4.2 Fuel Injection Control

The delay that is present in the lambda feedback-control limits the performance at sudden changes in, e.g., air-flow due to throttling. Therefore, it is necessary to combine the feedback path with feedforward paths. Together, these form the fuel-injection control-system. Sensors are used to measure air and coolant temperatures, pressures, air flow, throttle position, crankshaft position, and engine knock. Actuators are used to control fuel injection, airflow, and EGR. The fuel supply to the injectors is normally under constant, controlled, pressure. The injectors are controlled by their open-time (typically 1–10 ms), which is then proportional to the amount of injected fuel. The primary decision variables for the ECU are the engine intake air-mass flow, $\dot{m}_a$, and the rotation rate (rpm), N. The air mass inducted by a cylinder per cycle is then $m_a = \dot{m}_a/N$. From this, a basic injection opening duration is computed from a lookup-table, or map. The basic opening duration is then modified with corrective factors depending on engine temperature, rate of throttle opening, and lambda-sensor signal. Also these factors are computed from look-up tables.

The airflow into the engine may be determined directly with an airflow meter. Since these are quite complex and expensive* the indirect speed-density method is often used instead. Then, the engine speed, inlet-manifold absolute pressure, and air temperature are used to compute the inducted air mass. The manifold pressure is also used to measure engine load and may be measured by a cheap sensor. The mass of air inducted per

*A fact the author is sadly familiar with.

cycle is $m_a = \eta_v V_d \rho_a$, where V_d is the cylinder volume, η_v the volumetric efficiency, which varies with engine speed, and ρ_a the air density, which is a function of the inlet air temperature and pressure.

Idle-speed is controlled with a throttle bypass passage, or, if present, with an electronically controlled throttle, such that the air-mass may be adjusted to keep the engine-speed constant, also in presence of disturbances. PID-controllers are commonly used for this task.

4.3 Ignition Control

Maximum efficiency of the engine is obtained if the cylinder pressure-peak resulting from the combustion appears just after the piston starts to move down in the power stroke. The piston position is described with the crank-shaft angle, where 0 degrees, or top dead center, denotes upper position in the power stroke. Having the pressure peak at the desired position implies that the fuel must be ignited before the piston reaches the top dead center position. This is denoted the ignition-advance angle, which varies with engine speed, air/fuel ratio, and manifold pressure. The ignition advance angle is computed from a look-up table, the ignition map, using engine speed and manifold pressure. Other maps provide with correction factors for coolant temperature, throttle position and injection duration.

The ignition-advance angle is normally set as early as possible for efficiency reasons. If it is set too early the mixture may start to burn before the top dead center, which results in engine knock. This is detected with, e.g., a piezoelectric vibration sensor mounted in the engine block. The EMS monitors the knock sensor a few milliseconds after each spark and if knock is detected the ignition-advance angle is reduced. It is then slowly restored to its nominal value. In this way, the engine always operate close to the knock limit, for efficiency reasons.

4.4 Electronic Throttle Control

The traditional air-throttle was mechanically linked to the accelerator pedal. By pressing the pedal the driver directly determined the air-flow into the engine. With the electronically controlled throttle the accelerator pedal is equipped with position and possibly force sensors monitoring the drivers intentions. The throttle is equipped with a servo-motor and a throttle-angle sensor. The pedal-sensor signals are translated to a desired throttle-angle and a feedback controller positions the throttle at the setpoint. A production system is exemplified in [McKay *et al.*, 2000].

The Electronic Throttle Control (ETC), or Drive-by-Wire, gives flexibility in designing the feel of driving, as the accelerator pedal signals may be interpreted differently under different conditions. With ETC, throttle actuation may also be used by other control systems such as adaptive cruise control, traction control, idle speed control, and vehicle stability control.

4.5 Paper III

Lean-burn engines have reached the market in recent years and provide a means to reduce fuel consumption as well as CO and NOx emissions. Those engines are operated at high air-fuel ratios close to the combustion limits and, therefore, need very precise control. In these engines a new type of EGO sensor is used, the UEGO sensor, which gives a continuous signal over the operating range of air/fuel ratios, in contrast to the traditional relay-type EGO sensor.

The Gasoline Direct Injection (GDI) engine is a special type of lean-burn engine [Zhao *et al.*, 1997, Iwamoto *et al.*, 1997]. In GDI engines the fuel is injected directly into the cylinders, in contrast to traditional port fuel-injection. The GDI engine can operate in two different combustion modes. In the homogeneous combustion mode a homogeneous air/fuel charge is used together with stoichiometric combustion. In the stratified mode the mixture formation is such that the charge is rich only close to the spark plug and combustion is, therefore, made possible at very high air/fuel ratios.

To simplify the interfacing with other systems, engine commands are preferably expressed as desired engine torques. To comply with this, modern engine management systems are increasingly oriented towards torque-control architectures, as in [Gerhardt *et al.*, 1998]. The torque control is, in general, not realized with feedback from torque sensors because of the lack of low-cost solutions. Internal torque-production models are instead used. Much research is, however, devoted to searching for a low-cost torque-measurement method. Candidate methods are surveyed in [Turner and Austin, 2000].

Paper III deals with torque-control of a GDI engine. Engine-controllers today depend heavily on having accurate engine characteristics stored in look-up tables or maps. In the controller-scheme presented in Paper III, some of these maps are replaced by rough linear approximations and feedback from the engine torque. The result is a controller that is more robust to engine variations and disturbances. Optimal operating points with respect to fuel economy are determined on-line using extremum control.

Extremum Control

In [Kiencke, 1988] a self-optimizing control of fuel consumption is described, in which the optimum operating condition of the engine may be found by superimposing an orthogonal test function $\Delta\dot{m}_L = m_0 \cos \omega t$ (air modulation) to the intake air $\dot{m}_L$ and cross-correlating it with an output signal dn/dt proportional to engine torque. The stationary behavior of the engine may be characterized by a quadratic function of specific fuel consumption over air [Schweitzer *et al.*, 1966]. The result of the correlation contains the gradient of the the quadratic function as

$$\text{gradient} = \frac{1}{T_m} \int_{t-T_m}^{t} \Delta\dot{m}_L \left(t - (\alpha/\omega)\right) \left(\frac{dn}{dt}\right) dt \tag{4.1}$$

where T_m is the correlation time, and α the phase lag resulting from engine dynamics at the frequency ω. The gradient is regulated to zero by integral control. It is noted that the air may be modulated no faster than about 2 Hz. With a sufficiently large correlation time the convergence time goes up to about 30 s. Therefore, applications are considered to be limited to calibration of open-loop engine maps.

The novelties of the scheme proposed in Paper III are the combination of an optimizing controller with a torque-controller and the multi-variable setup where also the EGR flow is optimized. The torque controller rejects the disturbances introduced by the probing signals, or test functions, and render them invisible in the engine output. The control signal from the torque controller is used for correlation in the optimizing controller, instead of the torque signal. In the simulation study the convergence of the controller was faster than reported in [Kiencke, 1988], but in a real application the performance may very well decrease. Still, the convergence time is believed to be sufficiently short for optimizing engine operation during steady operation, such as highway driving with cruise-control. Extensions of the scheme to optimizing other variables such as ignition-timing may also be considered.

5

Systems Implementation

5.1 Systems Architectures

High-end cars of today are equipped with an advanced computer architecture consisting of sensors, actuators, computer nodes, and a communication network connecting these components. On this hardware infrastructure a complex distributed control system is hosted.

The computer nodes belonging to a vehicle-control system are commonly denoted as electronic control units (ECU). The ECUs consist of microcontrollers and I/O-interfaces for network, sensors, and actuators. Nodes for body electronics are in general simple with, typically, 8-bit microcontrollers with 100 byte RAM and 32 kbytes ROM. Chassis and powertrain electronics usually have more powerful nodes hosting 16 or 32-bit microcontrollers, with around 16 kbyte RAM and 512 kbyte ROM, and possibly floating-point capacity. The engine ECU is, normally, the most powerful microcontroller on the vehicle with 32-bit technology and floating-point capacity.

Computers were first introduced in cars on a large scale with the first generation of electronic engine control in the late 1970s. As the number of ECUs increased with the introduction of more computer-based control technology, the need for information sharing between these also increased. It was early realized that the issue of cabling was to become a major obstacle for further integration of ECUs, as growing wiring harnesses would be too complex and would add too much weight to the vehicle. A remedy to this was the introduction of multiplexed digital networks, enabling many signals to share one cable. The development of the controller area network (CAN) for automotive applications started already in 1983. In 1992 the first car with CAN appeared on the market and CAN is now the dominating standard for connecting ECUs [Robert Bosch GMBH, 1991].

CAN is a serial, asynchronous, CSMA/CA (Carrier Sense, Multiple Access with Collision Avoidance) protocol for multiplexed digital buses. The physical layer uses a pair of wires and comes in two main flavors: a high-speed (up to 1 Mbit/s) and a low-speed (up to 125 kbit/s) architecture. Each message type has a unique identifier that also translates to priority. Bus arbitration is synchronized so that all candidate sending nodes except the one with the highest priority message will cancel their transmission in each round. This means that the transmission timing of lower priority messages is stochastic.

The requirements on the communication network depend on the system functions connected to it. The SAE has identified three classes:

Class A Low speed networks (<10 kbit/second).

Class B High speed networks with no safety critical requirements (10 kbit/s to 125 kbit/s).

Class C High speed networks with stringent safety-critical requirements (125 kbit/s to 1 Mbit/s or greater).

Low-cost class A networks are, typically, used for networking body-control functions and diagnostic information with event-driven message transmission. Body-control response times are, in general, around 100–200 ms. Common protocols are low-speed CAN and Local Interconnect Network (LIN) [von der Wense, 2000]. Class B networks are used for all non safety-critical information-sharing between ECUs in powertrain and chassis controls, often with timing requirements around 1–20 ms. High-speed CAN is is the dominating protocol. Safety-critical X-by-wire technology and advanced chassis-control pose high demands on communication requirements, with respect to fault-tolerance and timing constraints. All fault scenarios must be accounted for with a safe, alternative operating mode and the bus must be guarded against faulty nodes. Low latency and jitter must be ensured if high-bandwidth control loops are to be closed over the network. Event-driven protocols such as standard CAN are not considered to meet these requirements. Time-triggered protocols that are based on synchronized nodes and scheduled message transmission give deterministic timing and are by many considered as the only feasible solution. New competing protocols aimed at class C networking are the TDMA (Time-Division, Multiple-Access) Time-Triggered Protocol (TTP/C) [Kopetz and Grunsteidl, 1993, Kopetz and Bauer, 2003] and FlexRay [Belschner *et al.*, 2002]. The latter providing for both event-triggered and time-triggered message transmission. There are also efforts to enhance CAN to a time-triggered protocol, TTCAN [Führer *et al.*, 2000, Leen and Heffernan, 2002].

Current road vehicles are equipped with several different networks for, e.g., body, powertrain, and chassis control. The networks are connected with gateways to enable information sharing. A subsystem connected to a network may also have a local sub-network for sensors and actuators. The overall system forms a quite complex, hierarchical, heterogeneous network, see Figure 5.1.

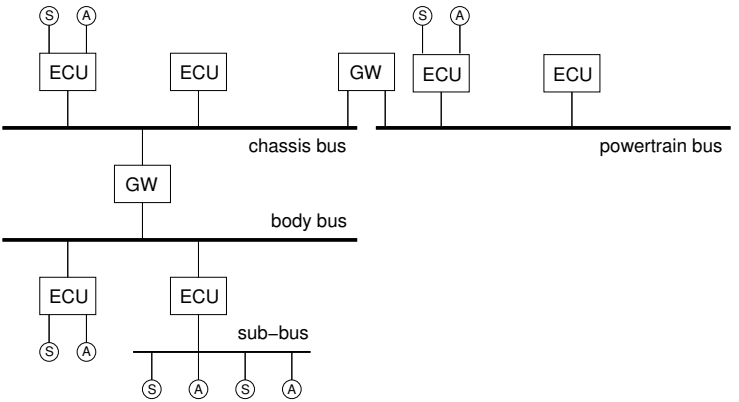

Figure 5.1 A subset of a typical vehicle-network configuration with ECU, gateway (GW), sensor, and actuator nodes.

Open Architectures

An obstacle on the path to integrated vehicle-control systems is that suppliers often deliver complete subsystems that are not based on open architectures. As a consequence, different subsystems may be incompatible due to, e.g., the use of different protocols and interfaces. This makes it difficult and costly for car manufacturers to integrate systems from different suppliers and decreases the flexibility in the total system design.

To address this problem there are efforts by the automotive industry to introduce standardized open architectures based on functional decomposition and common interfaces. A major goal is to obtain efficient architecture designs such that functionalities are configurable and scalable. A European initiative is offered by OSEK [OSE, 2000] (English "Open Systems and the Corresponding Interfaces for Automotive Electronics"). The aim is an industry standard for an open-ended architecture for distributed control-units in vehicles. Standards are defined for communication, network configuration management, operating system, and imple-

mentation language. A case-study on using open architectures in the design of a distributed brake-by-wire system is described in [Coelingh *et al.*, 2002, Kelling and Heck, 2002].

5.2 Dependability

Faults, Errors, and Failures

When discussing dependability of systems a distinction is made between *faults*, *errors*, and *failures* [Storey, 1996]. A fault is a defect within the regarded system. An error is a deviation from the required operation of the system or subsystem. A system failure occurs when the system fails to perform its required function.

In an automotive control-system faults may occur in the ECU hardware or software, or in the network. Faults are classified as permanent, intermittent, or transient. Transient faults are short-duration faults that are induced by, e.g., neutrons and alpha particles in semi-conductor devices, power supply and interconnect noise, electrostatic discharge, or electromagnetic interference (EMI). On certain highway passages in Europe, the engines of some vehicles have been shut off when their engine ECUs have been interfered by EMI from high-voltage lines beneath the roadway [Berger, 2002]. Intermittent faults are re-occurring short-term faults that occur due to marginal hardware or aging effects. Permanent faults, in general, have the same causes as intermittent faults, but reflect irreversible physical changes. In general, the development of integrated circuits with lower voltage and decreasing size leads to increasing noise sensitivity and strong reduction in reliability [Manzone, 2001]. As system complexity grows, an increasing number of faults are also the result of design-errors in hardware and software.

Errors can be classified as either data errors or timing errors. A data error occurs when the faulty component delivers data that is incorrect and a timing error occurs when data is delivered at an incorrect point in time.

System failure is defined with regard to system specifications. E.g., in a control system, failure may be defined as the exceeding of bounds on the control error, as the escape of the system state from a permitted region, or as loss of stability.

Computer-Architecture Fault-Tolerance

Faults in the computer hardware of a safety-critical automotive control system may lead to catastrophic failures, possibly with fatal outcome. Efforts to increase fault-tolerance is therefore a major part in the design

of such systems, see e.g. [Isermann *et al.*, 2002, Manzone, 2001, Kopetz, 1999]. Failure-rates must be kept very low because of the large number of vehicles in operation. Over 30 million cars were produced and 246 million registered in 1970. By 1997 these figures had risen to 56 million and 709 million, respectively, and the predictions for 2005 are more than 65 million cars produced and over 800 million registered [Powers and Nicastri, 2000]. Typical dependability requirements regard the mean time between failures (MTBF), fault detection and isolation, and avoidance of single point of failures.

Fault-tolerant distributed computer systems are traditionally based on techniques and paradigms from computer-engineering [Jalote, 1994]. In general, the detection of occurring faults is a central part. If fault-detection can be ensured, then various mechanisms for recovery can be activated at fault occurrence. These may regard error-handling, fault diagnosis, fault isolation, system reconfiguration, or system re-initialization. The effectiveness of fault-detection mechanisms are quantified by the coverage factor, which is the probability of detecting an occurring fault. Also, the use of redundancy is central and may cover replication of both hardware and software. A distinction is made between active redundancy, where all replicas are active in normal operation, and passive redundancy, where replicas are activated in case of a fault in the primary component. Components are preferably designed to be fail-silent. This means that a component stops producing data in case of a fault, so that propagation of the fault to other parts of the system is inhibited. Fail-silence also eliminates the need for voting mechanisms between replicas in an active-redundancy configuration, since data received from a fail-silent component will always be correct. Replica determinism is a property which ensures that the states of all correct replicas are consistent. It is based on methods to make the replicas reach agreement on current information. Network protocols with atomic broadcast services, such as TTP/C, ensure the delivery of all messages in the same order to all correct recipients and are often necessary in realizing replica determinism. Non-determinism may result from non-deterministic network protocols, the use of different sensor values, divergent software implementations, or the use of non-deterministic software constructs.

5.3 Paper IV

Methods and paradigms from computer engineering on construction of dependable distributed computer-systems are often disregarding properties of the application to be hosted in the system. It is, in general, assumed that all failure handling and avoidance is taken care of by the computer

system. Moreover, it is assumed that failure to fulfill the requirements at any instance leads to system failure. If the computer system hosts an implementation of a control system, this is not necessarily true. It may be possible for the control algorithms to take care of a failure in the computer system, transient or permanent, and continue to run the system in a safe, possibly restricted, mode. This reduces the number of faults that need to be managed by computer fault-tolerance mechanisms and makes it possible to relax the requirements on the computer system. This may then lead to a safer system with decreased complexity and cost.

The effects of timing errors on control systems have been investigated by many, see e.g., [Kim and Shin, 1994, Wittenmark *et al.*, 1995, Sanfridson, 2000]. The effect of transient data errors caused by EMI bursts on stability in control-system implementations has been investigated in [Kim *et al.*, 2000]. Catastrophic failures in a safety-critical system may occur before the system reaches instability, e.g., if some constraint on the control error is exceeded. Recent results [Cunha *et al.*, 2001, Vinter, 2001] show that many data errors will have a limited effect on control performance, i.e., control systems often have an inherent resilience or inertia to data errors. In [Askerdal *et al.*, 2002] suitable models of computer-faults for the analysis of control-level effects are presented, together with analysis methods for understanding the effect of data errors on system dependability, under the assumption of a linear time-invariant closed-loop system. Related results are presented in [Askerdal *et al.*, 2003].

In Paper IV this path is pursued further and a simple method to render a controller more robust to transient data-errors is presented. Transient faults have a temporary effect, but will in general leave the affected component in an inconsistent state. The proposed method automatically detects data-errors that result from occurring transient faults in the controller implementation and then restores the controller state. The idea is to compute bounds on the control signal, using known bounds on reference signals and disturbances, and then introduce an artificial control-signal limitation in combination with an observer based anti-windup scheme. When the limitation is reached, as a fault occurs, the controller states are forced back to feasible values by the observer dynamics.

5.4 Case Study

The publications [Claesson *et al.*, 2000, Sanfridsson *et al.*, 2000, Gäfvert *et al.*, 2000, Gäfvert, 2001] present a case-study on the architectural design of a brake system on a tractor-semitrailer combination vehicle, with certain additional functionality. The aim of the work is to analyze designs of a safety-critical distributed automotive control system, by applying and

combining methods within the areas of computer engineering and control theory. The brake system is a distributed safety-critical control system by nature and constitutes an excellent case study for this purpose. An increasing number of systems with similar requirements and constraints regarding performance, dependability, cost, etc. are found in the area of automotive engineering. This section presents a summary of the results. Details are found in the cited reports and in [Claesson, 2002].

The case study is based on the belief that there is a great potential in combining theory from the areas of computer engineering and control theory in the construction and design of dependable distributed control systems. The motivation for the work is the possibility of better, more effective system designs if the combination of the computer system and the control system is regarded at an early stage. This can be reached by ensuring that the computer system has certain properties from the control engineers point of view, and vice versa, by ensuring that the control system has certain properties from the computer engineers point of view. To successfully carry out codesign of the computer system, the communications network, and the control system it is necessary to specify interfaces between these systems in the design process and in the implementation. In the design process these interfaces consist of requirements and specifications, and in the implementation they may consist of certain functionality, such as software services or functions. It is necessary to specify the requirements in a detailed and systematic way. This can, for example, include what services the computer system and communications network should provide and what bandwidth requirements they should fulfill. For the control system this may be specifications that the control algorithms should be able to handle network delays of certain characteristics, transient failures such as lost sensor data for a certain time period, mode switches to safe limp-home modes, etc.

In the case-study specific focus is put on the choice of distribution level, or partitioning level, of the computer architecture and the control algorithms. Partitioning of the control algorithms is essentially a functional decomposition. Partitioning of the computer architecture and the control algorithms are orthogonal in the sense that they can be chosen independently and then be combined to form a complete system architecture. To explore different combinations, three conceptually different partitioning levels are regarded for the control algorithms and computer architecture respectively. These are detailed in the following sections. Combinations of these are then analyzed, resulting in a proposal on a system architecture. Techniques for partitioning and decentralization of real-time control systems are discussed in [Törngren and Wikander, 1996].

The Brake-System

The considered system is an electrical brake system (EBS) for a tractor-semitrailer vehicle, with functions for base braking, anti-lock braking (ABS), and active stability control (ASC). The ASC stabilizes the vehicle by applying unilateral braking actions when tire friction limits are reached during cornering, which prevents undesired and hazardous over- and understeering. In essence, it is a controller with state-feedback from the observed vehicle yaw-rate and side-slip, which is gain-scheduled on velocity [Kiencke and Nielsen, 2000]. The ASC computes command signals to dedicated brake-slip controllers for each wheel, which are also used for ABS and base-braking. A conceptual view of the brake system is shown in Figure 5.2. The algorithms compute signals for the pneumatic brake actuators, based on a number of sensor signals for wheel-speeds, yaw-rate, lateral and longitudinal acceleration.

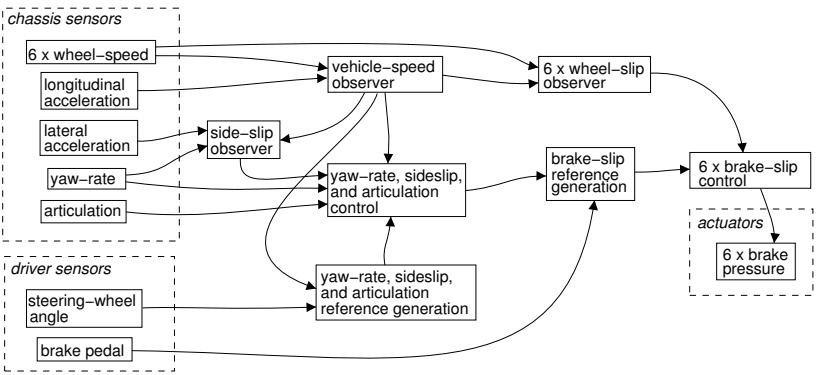

Figure 5.2 Conceptual view of the case-study brake-system.

Physical Partitioning

A distributed control system consists of a plant to control and several sensors, actuators, and controllers that are connected by means of a communications network. In a distributed control system the sensors and actuators are usually physically constrained to certain spatial positions. This may be denoted the *physical partitioning* of the system. For the brake system regarded in the case-study the physical distribution is illustrated in Figure 5.3.

Algorithmic Partitioning

On another level there is the *algorithmic partitioning*. This reflects the possible partitioning of the control algorithms into computational blocks

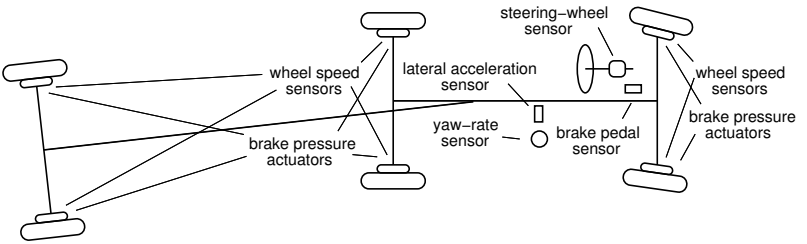

Figure 5.3 Physical distribution of sensors and actuators.

for, e.g., reference-signal generation, state observation, and control-signal computation. The diagram of Figure 5.2 indicates natural atomic building-blocks for partitioning of the brake-system.

Centralized Structure. One choice is to merge all atomic blocks into one central computational block, resulting in the structure of Figure 5.4. This is a centralized controller consisting of one monolithic algorithm,

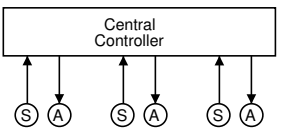

Figure 5.4 A centralized control-structure where "S" denotes sensors and "A" actuators.

that collects data from all sensors and computes outputs to all actuators. An advantage with this structure is the availability of all signals simultaneously. The lack of modularity, however, may make it difficult to modify the algorithms, to add new functionality, or to integrate the system with other co-existing systems.

Hierarchical Structure. Another choice is to partition the system into one coordinating top-level block and several auxiliary sub-blocks, resulting in the structure of Figure 5.5. This is an hierarchical partitioning that is natural in many control problems. Typically, the sub-blocks may include state observers or inner loops of cascade-control structures. In the brake-system application it is natural to have sub-blocks with dedicated brake-slip controllers for each wheel individually and a top-level block that coordinates braking actions by providing the brake-slip controllers with reference values.

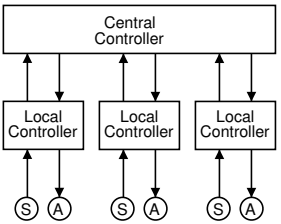

Figure 5.5 An hierarchical control-structure.

Peer Structure. There is also the possibility to partition the algorithms without any explicit hierarchy. Then, there is no block with the specific task of coordination and this is instead achieved by peer-to-peer information-sharing between the blocks, see Figure 5.6. For the brake-

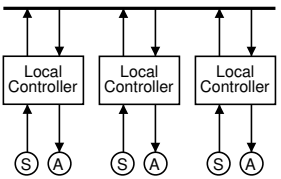

Figure 5.6 A peer control-structure.

system considered this is not a natural choice since the brake-slip control inherently is subordinated the base-braking and yaw-control functionality.

Computational Partitioning

On another level there is the *computational partitioning*, which determines how the computational resources are distributed. This corresponds to the partitioning and distribution of the computer and communications system.

Centralized Architecture. In the centralized architecture shown in Figure 5.7 all processing takes place in one single node. Sensors and actuators are connected to this node via a bus. This architecture has the advantage of requiring little hardware, although the single node may need to be powerful. This may keep down cost and maintenance. The simplicity of the system makes it less complex to manage with respect to dependability. Bandwidth requirements are quite easily verified.

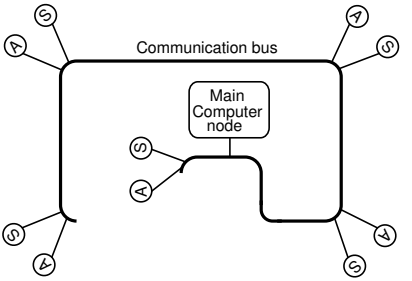

Figure 5.7 A centralized computer architecture.

Partially Distributed Architecture. In the partially distributed architecture of Figure 5.8 the sensor and actuator nodes are equipped with computing capacity. These nodes can be "smart" sensors and actuators, or hosts of local control loops. The decomposition into several units may simplify development and make the system more easy to test. Managing dependability is more complicated than on the centralized system.

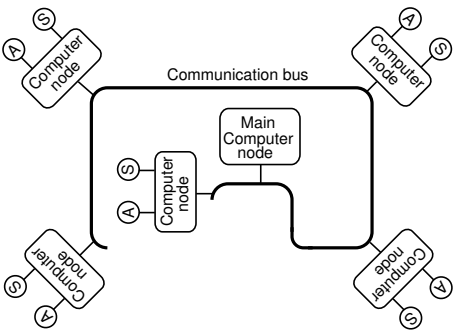

Figure 5.8 A partially distributed computer architecture.

Fully Distributed Architecture. In the fully distributed architecture of Figure 5.9 all nodes have the capacity to host complex computations. A substantial advantage with this architecture is the potential flexibility to add functionality in arbitrary nodes. While the partially distributed system is decomposed into small units of hardware, these are often dedicated to a particular functionality. The decomposition in the fully distributed system is completely realized on the functional level. Management of dependability may be quite complex, but the sensitivity to node failure may be smaller, since a peer node may take over functionality.

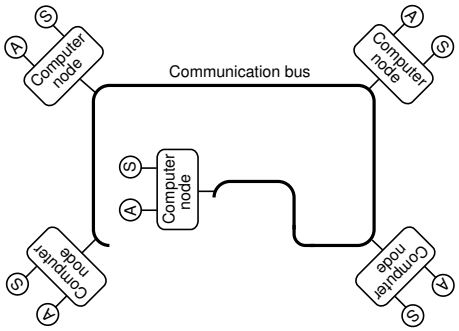

Figure 5.9 A fully distributed computer architecture.

System Architectures

A system architecture is obtained by mapping a control-system structure onto a computer architecture. The mapping determines where network delays are introduced in the algorithms. Note that the mapping may include replication of certain algorithm components into several nodes and that also this affects the location of network delays. In the case-study, centralized and hierarchical control structures were mapped onto centralized, partially distributed, and fully distributed computer architectures. The peer control-system architecture was not considered, since it was not natural to apply on the brake system. The investigated configurations are shown in Table 5.1. The mapping procedure involves a lot of freedom with respect to replication, etc., and not all variants make sense. Details on the chosen mappings are given in [Claesson *et al.*, 2000]. The C/C and H/C configurations show only node-internal differences, and therefore only C/C was considered. In these configurations all control algorithms execute in the only computer node. In the C/P architecture the distributed nodes may take over functionality if the central node fails. In the H/P configuration, the top-level algorithm execute in the central node and the local algorithms in the distributed nodes. This is a very intuitive configuration that is close in spirit to currently used architectures. In the C/F configuration all control algorithms are replicated in every node. This means a massive replication that may be exploited for increased dependability. For the same reason, the H/F architecture is interesting from a dependability point-of-view, as it is assumed that replicas of the top-level algorithm are executing in all nodes. Similar architectures have been suggested for flight-control systems. Due to large bandwidth requirements and complexity it is, however, considered as a less attractive solution for automotive systems.

Table 5.1 Abbreviations of system architectures. Configurations within parentheses are not exhaustively analyzed.

Control structure	Computer architecture		
	Centralized	Partially distributed	Fully distributed
Centralized	C/C	C/P	C/F
Hierarchical	(H/C)	H/P	(H/F)
Peer	(P/C)	(P/P)	(P/F)

Dependability. The full functionality of the brake system includes base braking, ABS and ASC. Faults that result in loss of ABS and ASC functionality will, generally, not result in immediate hazard. Only the base-braking is therefore considered to be a critical functionality that must not fail under any circumstances. To ensure critical functionality in the case of single faults, certain components are duplicated. Two cases are therefore regarded for each computer architecture: a simplex configuration with no redundancy and a duplex architecture with active redundancy of critical components.

Dependability measures are computed from assumptions on which components are used in a particular computer architecture, in combination with component failure-rates obtained from tabular data on electronic equipment. Probabilities of system failure for the considered computer architectures with simplex and duplex configurations are shown in Table 5.2. Probabilities of loss of any or critical functionality for the considered computer architectures are shown. In the calculations the inherent actuator-redundancy of the brake system is taken into account. A minimum brake functionality requires one operational actuator on each side of the vehicle. The absolute values of the failure probabilities have no significant meaning. The relative values, however, indicate the relative dependability of the different architectures. It is clear that a duplex configuration is preferred to obtain high dependability. The choice of C, P or F architecture is of less importance. The differences between the P and F architectures are due to assumptions on simpler hardware in the distributed nodes of the P architecture. Observe that the coverage factor has a very large influence for the duplex configurations. That means, it is critical to have effective fault-detection mechanisms.

Bandwidth Requirements. Which signals that are sent over the data-bus depend on the system configuration. Bandwidth requirements are computed from signal data-formats and periodicities. Results for the considered configurations are shown in Table 5.3. Replica management for

Table 5.2 Probabilities of loss of any or critical functionality for the considered computer architectures with simplex and duplex configurations. Note that the coverage C has large influence on the dependability for the duplex configurations.

Functionality	Computer architecture		
	C	P	F
Any (simplex)	3.3×10^{-4}	3.5×10^{-4}	3.5×10^{-4}
Critical (simplex)	4.3×10^{-5}	2.7×10^{-5}	3.0×10^{-5}
Critical (duplex, $C = 0.990$)	3.7×10^{-7}	2.2×10^{-7}	2.7×10^{-7}
Critical (duplex, $C = 0.995$)	1.8×10^{-7}	1.1×10^{-7}	1.4×10^{-7}
Critical (duplex, $C = 0.999$)	3.8×10^{-8}	2.2×10^{-8}	2.7×10^{-8}

Table 5.3 Bandwidth requirements [bit/s] for the considered system architectures.

	System architecture			
	C/C	C/P	C/F	H/P
Simplex	23,280	23,280	8,720	23,280
Duplex	57,680	29,440	15,680	30,240

the duplex systems requires significant bandwidth, which explains the numbers seen in the table. The duplex C/C configuration has the largest bandwidth requirement. This is because all sensor and actuator data is sent on the bus with high frequency. The smallest bandwidth by far is obtained with the C/F configuration, where many signals are kept internal in nodes with local control loops.

Conclusions

The duplex system with the smallest bandwidth requirements is C/F. With replication of the full control-algorithm in all nodes the reconfiguration after a fault is much simplified. In general, dedicated central nodes should be avoided if possible, as such architectures may require complicated re-configurations in case of a fault. This is in favor of decentralized architectures with replication of centralized algorithmic components. The highest reliability was achieved with the H/P architecture, which is the configuration most similar to current systems. A H/F configuration without replication is very similar to H/P, except for having more powerful nodes in the former. This configuration offers great flexibility in system design as new functionality may be added in any node. This implies a clear separation between computer architecture and control algorithms

and requires standardized interfaces and protocols, much in the spirit of the open architectures currently considered by the industry.

6

Concluding Remarks

A general conclusion from the work on this thesis is that areas that seem unrelated at first sight become clearly related when being part of the same system. With knowledge only on isolated parts or topics, it is impossible to gain complete understanding of a system as a whole. In automotive applications this is very clear. For several decades the road vehicle was mainly a mechanical construction. In recent years, computer technology has been introduced as a key component and advanced control algorithms are used to modify the dynamic behavior. It is a very satisfying experience when new insights on the system level are derived by combining knowledge from these, and other, areas.

Paper I

A useful observation from this work is that vehicle models with sound structures and of intermediate detail level can be straightforwardly constructed with general computer-algebra tools and simulation software. This is an alternative to the models of similar detail level that are offered commercially. An advantage is that insight into the modeling details may help the user to better understand the possibilities and limitations of their model. For the purpose of trying out new ideas on chassis control the approach is ideal, since actuating forces and moments may be applied at will. The modular formulation makes it very easy to change vehicle configuration, such as adding axles or changing suspension characteristics, and the graphical interface of Simulink offers an attractive view. The nonlinear description of the chassis dynamics makes the model applicable in a wide range of operating conditions, also at large articulation angles. The performance and accuracy turned out to be satisfactory, even if optimization of the final implementation seems necessary to obtain realtime simulation speed.

Extensions and Future Work

For hardware-in-the-loop simulations an implementation completely in C, together with a suitable solver library, would maybe be the best solution. Another improvement in the direction of better efficiency would be to identify and eliminate terms that have little influence from the equations of motion.

There are many natural extensions to a model of this type. Virtually any subsystem of interest can be added. Detailed models of the steering subsystem and of the brake actuators would be especially useful. The presented model has mass-less axles. Attempts to include unsprung masses has been made already, but did not result in remarkably different results. It was noted that the increased complexity was still manageable with the modeling method used. Another addition to consider is the effects of vertical tire deflection, since this contributes significantly to the roll angle at load transfer.

Paper II

In the survey of existing tire models aimed at vehicle-handling simulations it was noted that the area is sparsely populated with accurate, or at least more advanced, models that are simple enough to be implemented by an end user. Many models were presented in the 1970s, or earlier, at a time when computer resources were scarce. The expressions involved are often overly simplified and have only rough physical foundation. In the eyes of the authors this is a niche where the proposed model fits well. The approach to base the model on the separation of forces from the adhesion and sliding regions turned out to give new intuitive and physically motivated ways to interpret and use the empirical data. As one result of this, the model includes effects of velocity dependence, even if this is not explicitly included in the used pure-slip models. Velocity dependence is a rare feature in models of similar size. Another very useful property of the model is that no additional parameters need to be introduced, which require calibration.

Extensions and Future Work

The quantitative validation that is shown supports the model, but is based on a limited data-set. It is clearly of interest to do a more exhaustive comparison with empirical data. However, it is not easy to find measurement data of good quality for all conditions of interest.

The proposed model may be viewed as a first step, as there is a natural series of refinements that can be made. Effects of camber are necessary to

include to make the model usable in a wide range of applications. This is expected to be straightforward, although a bit tedious. The model is based on the assumption of a rigid carcass, with some limited compensation for carcass deformation. Work is ongoing to include effects of carcass flexibility more completely, with the aim of even more accurate description of the lateral force and aligning moment. The proposed model describes steady-state behavior. In a further extension the carcass deformation may be described dynamically to include also transient behavior. This is of interest for vehicle simulations with rapid maneuvering. Among other issues to consider is the effect of normal-load dependence on combined-slip conditions. The model currently includes this implicitly, by means of the pure-slip models. Another considered feature is also the description of effects of water films on the road surface.

Paper III

Calibration of engine maps is a big issue in the design and production of engine controllers. The results indicate that control schemes that are more based on feedback than current approaches are less dependent on accurate maps. The bottom-up approach in the control design enabled the use of simple and mostly linear techniques. This is encouraging, and supports the general principle of always trying the simple first. The successful application of extremum control implies that this old idea may find application in the control of modern engines. The proposed method to combine the extremum controller with a controller that rejects the introduced probing-disturbances would then be a natural component.

Extensions and Future Work

In the average car on the street, the engine maps are rarely very accurate and it is likely that realtime optimization can give noticeable improvements. It seems particularly useful in combination with cruise-control systems in highway driving at constant operating conditions. A natural extension is to optimize also other variables, and to use another signal than the torque for performance monitoring. The method could be used to update the current engine maps in a vehicle, by starting an optimization from the old values. Another use would be for engine-map calibration in production sites. Then, variables that are not possible to measure during normal operation can be used and optimized, such as emission measures. Although not found in any references, it would be surprising if this was not tried already. Experiments on a real engine would be very interesting.

Paper IV

Currents trends in automotive computer-architectures and the semi-conductor industry are somewhat conflicting, as the dependability requirements on electronics increases in the former, while the latter delivers devices with decreasing reliability. Even if the fault-tolerance of systems may be kept on acceptable levels using redundancy, etc., this will, in general, induce higher cost and increased complexity. Results that enable the use of simple, cheap devices without jeopardizing safety are therefore welcome contributions and the method presented in this paper may be part of such solutions.

It is gratifying when old ideas may be re-used in new contexts. In this work the well known observer-based anti-windup scheme found a new application together with artificial signal limits. The simple idea that is presented effectively reduces the negative effects of occurring transient faults on control performance. It is also noted that results regarding finite-precision numerics in signal-processing are related to the effects of transient data errors.

Extensions and Future Work

The results extend straightforwardly also for rate-limited signals. An interesting path to pursue would be to determine the optimal realizations with respect to transient errors, as described in the paper, and to compare them with the optimal realizations for round-off noise.

7

Bibliography

Askerdal, Ö., M. Gäfvert, M. Hiller, and N. Suri (2002): "A control theory approach for analyzing the effects of data errors in safety-critical control systems." In *Proceedings of the Pacific Rim International Symposium on Dependable Computing*.

Askerdal, Ö., M. Gäfvert, M. Hiller, and N. Suri (2003): "Analyzing the impacts of data errors in safety-critical control systems." *IEICE Transactions on Information and Systems*, **E86-D:12**. Special issue on Dependable Computing. Invited paper. Submitted.

Bakker, E., L. Nyborg, and H. B. Pacejka (1987): "Tyre modelling for use in vehicle dynamics studies." SAE Techical Paper 870421.

Bakker, E., H. B. Pacejka, and L. Lidner (1989): "A new tire model with an application in vehicle dynamics studies." SAE Technical Paper 890007.

Bauer, H., Ed. (1996): *Automotive Handbook*. Robert Bosch GmbH.

Bayle, P., J. F. Forissier, and S. Lafon (1993): "A new tire model for vehicle dynamics simulations." *Automotive Technology International*, pp. 193–198.

Belschner, R., J. Berwanger, C. Ebner, H. Eisele, S. Fluhrer, T. Forest, T. Führer, F. Hartwich, B. Hedenetz, R. Hugel, A. Knapp, J. Krammer, A. Millsap, B. Müller, M. Peller, and A. Schedl (2002): *FlexRay: Requirements Specification Version 2.02*. BMW AG, DaimlerChrysler AG, Robert Bosch GmbH, General Motors/Opel AG.

Berger, I. (2002): "Can you trust your car?" *IEEE Spectrum*, April.

Bliman, P.-A., T. Bonald, and M. Sorine (1995): "Hysteresis operators and tyre friction models. Application to vehicle dynamic simulation." In *Proceedings of ICIAM 95, Hamburg, Germany*.

Brach, R. M. and R. M. Brach (2000): "Modeling combined braking and steering tire forces." SAE Technical Paper 2000-01-0357.

Brennan, S. and A. Alleyne (2001): "The use of a scale vehicle testbed for controller design and evaluation." *IEEE Control Systems Magazine*, **21:3**, pp. 15–26.

Canudas De Wit, C. and P. Tsiotras (1998): "Dynamic tire friction model for vehicle traction control." In *Proceedings of the 1998 American Control Conference*. Philadelphia.

Chowanietz, E. (1995): "Automobile electronics in the 1990s. Part 1: Powertrain electronics." *Electronics & Communication Engineering Journal*, **7:1**, pp. 5–10.

Claesson, V. (2002): *Efficient and Reliable Communication in Distributed Embedded Systems*. PhD thesis, Chalmers University of Technology.

Claesson, V., M. Gäfvert, and M. Sanfridsson (2000): "Proposal for a distributed computer control system in heavy-duty trucks." Technical Report 00-16. Computer Engineering, Chalmers University of Technology, Sweden.

Coelingh, E., P. Chaumette, and M. Andersson (2002): "Open-interface definitions for automotive systems." SAE Technical Paper 2002-01-0267.

Cunha, J. C., R. Maia, M. Z. Rela, and J. G. Silva (2001): "A study of failure models in feedback control systems." In *Proceedings of the IEEE International Conference on Dependable Systems and Networks*, pp. 314–323.

Dahlberg, E. (2000): *Commercial Vehicle Stability — Focusing on Rollover*. PhD thesis, Department of Vehicle Engineering, Royal institute of Technology, KTH, Sweden.

Deur, J. (2002): "A brush-type dynamic tire friction model for non-uniform normal pressure distribution." In *Proceedings of the 15th Triennial IFAC World Congress*. Barcelona, Spain.

Edlund, S. (1991): "Tyre models: Subreport -91." Technical Report. Volvo Truck Corporation. Confidential.

Führer, T., B. Müller, W. Dieterle, F. Hartwich, R. Hugel, and M. Walther (2000): "Time triggerd communication on CAN (time-triggered CAN — TTCAN." In *Seventh International CAN Conference (ICC)*. Amsterdam, Netherlands.

Gäfvert, M. (2001): “Studies on yaw-control of heavy-duty trucks using unilateral braking.” Technical Report ISRN LUTFD2/TFRT--7598--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gäfvert, M., M. Sanfridsson, and V. Claesson (2000): “Truck model for yaw and roll dynamics control.” Technical Report ISRN LUTFD2/TFRT--7588--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gäfvert, M. and J. Svendenius (2003): “Construction of novel semi-empirical tire models for combined braking and cornering.” Technical Report ISRN LUTFD2/TFRT--7606--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gerhardt, J., H. Hönninger, and H. Bischof (1998): “A new approach to functional and software structure for engine management systems - BOSCH ME7.” SAE Technical Paper 980801.

Gim, G. and P. Nikravesh (1991): “An analytical model of pneumatic tyres for vehicle dynamics simulations. part 2: Comprehensive slips.” *International Journal of Vehicle Design*, **12:1**, pp. 19–39.

Hac, A. (2002): “Influence of active chassis systems on vehicle propensity to maneouver-induced rollovers.” SAE Techical Paper 2002-01-0967.

Hecker, F., S. Hummel, O. Jundt, K.-D. Leimback, I. Faye, and H. Schramm (1997): “Vehicle dynamics control for commercial vehicles.” SAE Technical Paper 973284.

Holdmann, P., P. Köhn, and J. Holtschulze (1999): “Dynamic tyre properties under combined slip situations in test and simulation.” In *EAEC Barcelona 1999, European Automotive Congress*, pp. 817–822.

Isermann, R., R. Schwartz, and S. Stölzl (2002): “Fault-tolerant drive-by-wire systems.” *IEEE Control Systems Magazine*, **22:5**, pp. 64–81.

Iwamoto, Y., K. Noma, O. Nakayama, T. Yamauchi, and H. Ando (1997): “Development of gasoline direct injection engine.” SAE Technical Paper 970541.

Jalote, P. (1994): *Fault Tolerance in Distributed Systems*. Prentice Hall PTR.

Kelling, N. A. and W. Heck (2002): “The BRAKE project — centralized versus distributet redundancy for brake-by-wire systems.” SAE Technical Paper 2002-01-0266.

Kiencke, U. (1988): "A view of automotive control systems." *IEEE Control Systems Magazine*, **8:4**, pp. 11–19.

Kiencke, U. and L. Nielsen (2000): *Automotive Control Systems: For Engine, Driveline, and Vehicle*. Springer-Verlag.

Kim, H. and K. G. Shin (1994): "On the maximum feedback delay in a linear/nonlinear control system with input disturbances caused by controller-computer failures." *IEEE Transactions on Control Systems Technology*, **2:2**, pp. 110–122.

Kim, H., A. L. White, and K. G. Shin (2000): "Effects of electromagnetical interference on controller-computer upsets and system stability." *IEEE Transactions on Control Systems Technology*, **8:2**, pp. 351–357.

Kopetz, H. (1999): "Automotive electronics." In *Proceedings of the 11th Euromicro Conference on Real-Time Systems*, pp. 132–140.

Kopetz, H. and G. Bauer (2003): "The time-triggered architecture." *Proceedings of the IEEE*, **91:1**, pp. 112–126.

Kopetz, H. and G. Grunsteidl (1993): "TTP — a time-triggered protocol for fault-tolerant real-time systems." In *The Twenty-Third International Symposium on Fault-Tolerant Computing*, pp. 524–533. IEEE.

Leen, G. and D. Heffernan (2002): "TTCAN: a new time-triggered controller area network." *Microprocessors and Microsystems*, **26:2**, pp. 77–94.

Manzone, A. (2001): "Fault tolerant automotive systems: an overview." In *Proceedings of the Seventh IEEE International On-Line Testing Workshop*, pp. 117–121.

Maurice, J. (2000): *Short wavelength and dynamic tyre behaviour under lateral and combined slip conditions*. PhD thesis, TU Delft.

McKay, D., G. Nichols, and B. Schreurs (2000): "Delphi electronic throttle control systems for model year 2000; driver features, system security, and OEM benefits. ETC for the mass market." SAE Technical Paper 2000-01-0556.

Nguyen, P. K. and E. R. Case (1975): "Tire friction models and their effect on simulated vehicle dynamics." In *Proceedings of a Symposium on Commercial Vehicle Braking and Handling*, number UM-HSRI-PF-75-6, pp. 245–312.

Nicholas, V. T. and T. Comstock (1972): "Predicting directional behavior of tractor semitrailers when wheel anti-skid brake systems are used." In *ASME Winter Annual Meeting*, number Paper No. 72–WA/Aut-16.

Nordström, O. (1993): “The VTI flat bed tire test facility - a new tool for testing commercial tire characteristics.” SAE Techical Paper 933006.

OSEK (2000): *OSEK/ VDX Binding Specification Version 1.1*.

Pacejka, H. B. (1988): “Modelling of the pneumatic tyre and its impact on vehicle dynamic behavior.” Technical Report i72B. Technische Universiteit Delft.

Pacejka, H. B. (2002): *Tyre and Vehicle Dynamics*. Butterworth-Heinemann.

Palkovics, L. and A. Fries (2001): “Intelligent electronic control systems in commercial vehicles for enhanced traffic safety.” *Vehicle System Dynamics*, **35:4–5**, pp. 227–289.

Palkovics, L., A. Semsey, and E. Gerum (1999): “Roll-over prevention system for commercial vehicles — additional sensorless function of the electronic brake system.” *Vehicle System Dynamics*, **32**, pp. 285–297.

Pottinger, M. G., J. E. McIntyre, A. J. Kempainen, and W. Pelz (1993): “Truck tire force and moment in cornering - braking - driving on ice, snow, and dry surface.” SAE Techical Paper 933006.

Powers, W. F. and P. R. Nicastri (2000): “Automotive vehicle control challenges in the 21st century.” *Control Engineering Practice*, **8:6**, pp. 605–618.

Robert Bosch GMBH (1991): “CAN Specification Version 2.0.”

SAE Recommended Practice J670e (1976): “Vehicle dynamics terminology.”

Sampson, D. J. M. (2000): *Active Roll Control of Articulated Heavy Vehicles*. PhD thesis, Churchill College, Cambridge University Engineering Department.

Sanfridson, M. (2000): “Timing problems in distributed control.” Technical Report Licentiate thesis ISRN KTH/MMK--00/14--SE. Mechatronics Lab, Department of Machine Design, Royal institute of Technology, KTH, Sweden.

Sanfridsson, M., V. Claesson, and M. Gäfvert (2000): “Investigation and requirements of a computer control system in a heavy-duty truck.” Technical Report KTH/MMK/R--00/5--SE. Mechatronics Lab, Royal Institut, Stockholm, Sweden.

Schuring, D. J., W. Pelz, and M. G. Pottinger (1996): “A model for combined tire cornering and braking forces.” In *Investigations and Analysis in Vehicle Dynamics and Simulation*, pp. 61–83. SAE International. SAE Technical Paper 960180.

Schweitzer, P. H., C. Voltz, and F. Deluca (1966): "Control system to optimize engine power." SAE Technical Paper 660022.

Segel, L. (1984): "The mechanics of heavy-duty trucks and truck combinations." In *Engineering Summer Conferences*. University of Michigan, Ann Arbor.

Storey, N. (1996): *Safety-Critical Computer Systems*. Prentice Hall.

Törngren, M. and J. Wikander (1996): "A decentralization methodology for real-time control applications." *Control Engineering Practice*, **4:2**, pp. 219–228.

Turner, J. D. and L. Austin (2000): "Sensors for automotive telematics." *Measurement Science and Technology*, **11:2**, pp. R58–R79.

van Zanten, A., W. Ruf, and A. Lutz (1989): "Measurement and simulation of transient tire forces." SAE Techical Paper 890640.

Vinter, J. (2001): "Reducing critical failures for control algorithms using executable assertions and best effort recovery." In *Proceedings of the IEEE International Conference on Dependable Systems and Networks*, pp. 347–356.

von der Wense, H.-C. (2000): "Introduction to local interconnect network." SAE Technical Paper 2000-01-0153.

Wittenmark, B., J. Nilsson, and M. Törngren (1995): "Timing problems in real-time control systems: Problem formulation." In *Proceedings of the American Control Conference*, pp. 2000–2004.

Wong, J. Y. (2001): *Theory of Ground Vehicles*, 3 edition. John Wiley & Sons.

Zegelaar, P. W. A. (1998): *The dynamic response of tyres to brake torque variations and road unevennesses*. PhD thesis, TU Delft.

Zhao, F.-Q., M.-C. Lai, and D. L. Harrington (1997): "A review of mixture preparation and combustion control strategies for spark-ignited direct-injection gasoline engines." SAE Technical Paper 970627.

Paper I

A 9-DOF Tractor-Semitrailer Dynamic Handling Model for Advanced Chassis Control Studies

Magnus Gäfvert Olof Lindgärde

Abstract

This paper describes a flexible and modular 9-degrees-of-freedom nonlinear dynamic handling model for a tractor-semitrailer combination vehicle. The equations of motion are derived from the fundamental equations of dynamics in Euler's formulation, with the use of general computer-algebra software. The primary aim of the model is simulation of handling scenarios with active yaw control, using unilateral braking. However, it may also prove useful in other areas of tractor-semitrailer handling analysis or hardware-in-the-loop simulations. The model is formulated as a state-space model that may be implemented in standard simulation environments. A Simulink implementation is presented, and simulation results are compared with experiments to validate the model.

Figure 1 A Volvo FH12 tractor-semitrailer vehicle on the Öresund bridge. *(Courtesy of Volvo Truck Corporation.)*

1. Introduction

In the 1960s and 1970s there was an interest in handling-oriented models of tractor-semitrailer vehicles. The focus was then open-loop stability, and in particular jack-knifing behavior of the vehicles. In recent work the target has shifted to analyzing the handling of vehicles under closed-loop control with systems such as active yaw control. Models for simulation of tractor-semitrailer vehicles vary in a wide range of complexity from large multi-body system models with hundreds of degrees-of-freedom (DOFs) [Anderson *et al.*, 2001] to small 3-DOF bicycle models [Ellis, 1994].

In handling simulations the planar motions of the vehicle are of primary interest. These lateral, longitudinal, and yaw motions are driven mainly by the tire-road contact forces, which depend on the planar motion and the vertical tire load forces. The vertical forces vary with load transfer that results from inertial forces generated by the acceleration of vehicle masses. For commercial heavy vehicles this load transfer is large because of the high location of the center of mass (CM). A simple model to use for handling simulations would be a 4-DOF model: longitudinal, lateral, tractor yaw, and semitrailer yaw motion. For steady-state modeling

the load transfer may be described by e.g. quasi-static balance equations, or by effective tire characteristics [Pacejka, 1988]. For modeling of transient behavior it is necessary to introduce additional degrees of freedoms. On real vehicles the suspension systems have great influence on the load transfer. They introduce a phase lag between the planar motion of the vehicle axles, or unsprung bodies, and the vertical tire-road forces. To capture this phenomenon correctly it is necessary to use a model with sprung masses, that are connected to the unsprung body-parts by the suspension system. Combinations of heave, roll, and pitch freedoms may be used to describe the motion of the sprung masses.

Early work on 4-DOF planar models where load-transfer is modeled by an additional roll freedom or by quasi-static balance was performed by Ellis [Ellis, 1969, Ellis, 1988, Ellis, 1994] and Leucht [Leucht, 1970]. Mikulcik [Mikulcik, 1968, Mikulcik, 1971] presents an 8-DOF tractor-semitrailer model with longitudinal, lateral, heave, and roll freedoms for the tractor-semitrailer, and separate yaw and pitch freedoms for the tractor and semitrailer, respectively. A Fortran implementation of this model is also given. The model is used for analysis of jack-knifing behavior. In a recent work [Chen and Tomizuka, 1995, Chen and Tomizuka, 2000] a 5-DOF tractor-semitrailer model is derived using Lagrange methods. In this model a common roll freedom for the tractor and the semitrailer is introduced. The model is intended for the study of lateral dynamics control in the context of research on Automated Highway Systems. The model is evaluated with experimental data. In [Ranganathan and Aia, 1995] a linear 3-DOF model of a tractor-semitrailer at constant speed is derived and implemented in Simulink. The objective is yaw-stability analysis simulations. There are also recent intermediate DOF handling models like the commercial TruckSim software based on the AutoSim package [Sayers, 1992]. A drawback with those is that they are targeted at engineers rather than researchers. They are very easy to use when working with conventional vehicles, but have less support for unconventional designs that may be of interest in a research context. Another drawback with these tools is the closed environment they constitute. The models may be extended with equations expressed in Lisp-like language, but in an environment where many experimental designs are to be modeled it may be better to work with custom simulation models. Even though it may be tedious to work explicitly with the actual equations of a model, they provide a lot of insight into the model and its behavior.

In the present work the choice made was to include all the heave, roll, and pitch motions of the sprung masses, thus introducing five additional DOFs (the heave freedom is common for the tractor and the semitrailer). The reasons for this are that inclusion of pitch motion results in a more accurate description of longitudinal load transfer during braking and accel-

eration, and that the methods used for deriving the equations of motion do not become much more difficult to use with these extra freedoms. The modelling methodology relies heavily on computer-algebra software to handle large analytical expressions and to generate simulation code. There are several reasons for deriving a new model: Many previous models were derived by hand in the 1960s and 1970s and were, therefore, often subject to approximations that are not necessary with computer-algebra based methods available today. Furthermore, some equations were too large to solve analytically, and were thus left to numerical solution by iterative methods in the model implementations. This increases the computational complexity of the model, and may severely slow down simulations. These early models are also published as very large equations that are error-prone to implement in a simulation environment. Still these models provide good sources of knowledge when deriving a new model. The present model is also targeted at future real-time simulation applications. A possible use would be in hardware-in-the-loop simulations for evaluating various control equipment. This means that algebraic relations requiring iterative solutions during simulation have been avoided.

The paper is organized as follows: Section 2 describes the configuration of the vehicle. Section 3 treats the motion of the vehicle parts. In Section 4 internal and external forces are introduced, and the equations of motion are stated. The used tire model is described in Section 5. Section 6 describes the implementation of the model equations in a simulation environment. Results from simulations and validation of the model are then shown in Section 7. Finally, some concluding remarks are given in Section 8.

2. Vehicle configuration

The configuration of a model for the type of vehicle shown in Figure 1 is illustrated in Figure 2. The model consists of two sprung inertial bodies, connected by the hitch. The unsprung bodies (axles) are assumed to be massless. The hitch introduces algebraic constraints on the relative motion of the tractor and the semitrailer. These algebraic relations are used to reduce the number of freedoms in the final equations of motion. Suspension characteristics at the wheel corners are included, as well as torsional stiffness of the tractor chassis frame.

In the presentation below some notations are introduced. Vector variables are denoted by a bar. Vector representations with respect to a particular reference system are indicated by a superscript on the vector variable. Vectors are then represented as 3×1-matrices holding the components of the vector with respect to the reference system indicated by the super-

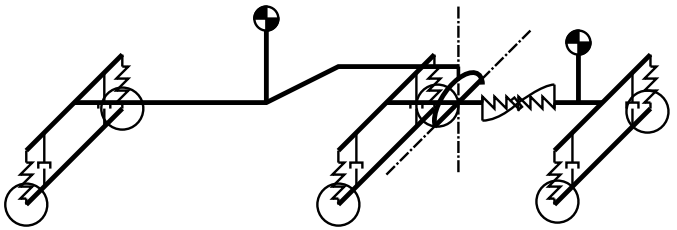

Figure 2 Vehicle configuration. The joint between the tractor and semitrailer is marked by an arc-shaped line representing the kinematic and kinetic constraints described in the text.

script. The base vectors of the reference system S_i represented in S_i are denoted by $\bar{e}_x^i = (1,0,0)$, $\bar{e}_y^i = (0,1,0)$ and $\bar{e}_z^i = (0,0,1)$. Coordinate systems are oriented according to the SAE standard, with x pointing forward, y to the right, and z downwards, with respect to the vehicle forward direction. Time derivatives of a vector with respect to a certain reference system are indicated by a subscript on the differential operator as $(d/dt)_i \bar{q}^j$. The shorthand notation of $\dot{\bar{q}}^j$ is used if $i = j$, i.e. the time derivative of $\bar{q}^j$ may be expressed as the time derivative of the components of $\bar{q}^j$. To help the reader, a table of notations is included in Appendix A. The analytical mechanics background is described in detail in [Lidström and Uhlhorn, 1984, Fowles and Cassiday, 1993].

The vehicle is divided into six separate parts, as shown in Figure 3: the tractor sprung body $\mathcal{B}_{ts}$, the semitrailer sprung body $\mathcal{B}_{ss}$, the tractor front axle unsprung body $\mathcal{B}_{af}$, the tractor rear axle(s) unsprung body $\mathcal{B}_{ar}$, the semitrailer axle(s) unsprung body $\mathcal{B}_{as}$, and the hitch* body or rear tractor part, $\mathcal{B}_h$. The main reference system is the coordinate system S_{to}. It is aligned with the tractor, but it neither rolls nor pitches. It is fixed with respect to $\mathcal{B}_{af}$ and $\mathcal{B}_{ar}$. The reference system S_{so} is similarly aligned with the semitrailer, and is fixed with respect to $\mathcal{B}_{as}$. The reference systems S_{ts}, S_{ss}, and S_h are attached to the corresponding bodies. The reference system S_* is the earth-fixed inertial reference system. The origins of S_{to}, S_{so}, S_{ts}, S_{ss} and S_h coincide and are located at the hitch, which is denoted O. This point on the vehicle has the unique property of being common for the tractor and semitrailer, which simplifies reduction of the equations of motion using the kinematic constraints.

The reference systems S_{to} and S_{so} describe yaw rotation and translational longitudinal, lateral, and vertical motion with respect to S_*. The angular velocities are $\bar{\omega}_{to}^{to} = (0,0,r)^T$ and $\bar{\omega}_{so}^{so} = (0,0,r')^T$, respectively.

*Also commonly denoted as "fifth wheel" or "king-pin".

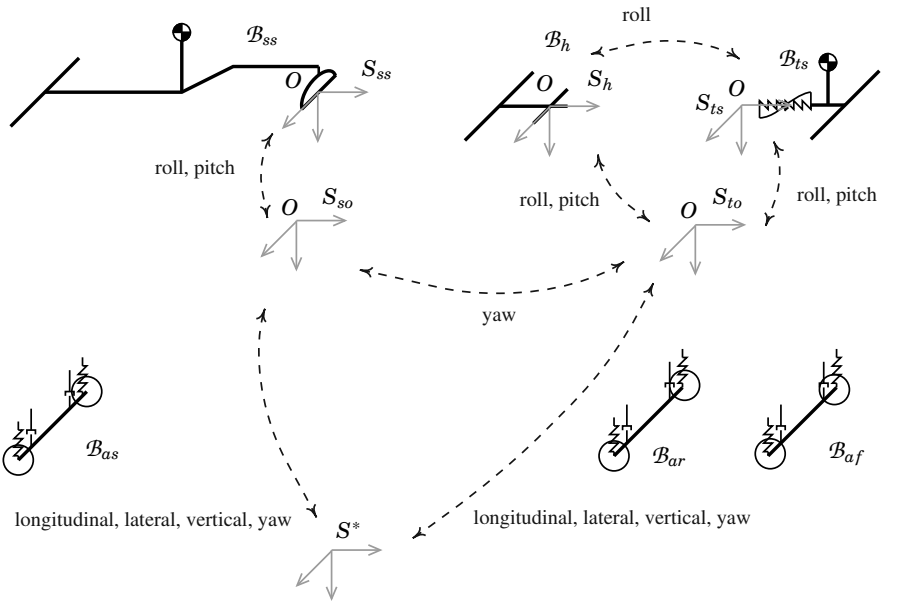

Figure 3 Vehicle parts with attached reference systems. The relative motions between reference systems are indicated by dashed lines and labels.

The velocity of O is $\bar{v}_O^{to} = (U, V, W)$. The angle between S_{to} and S_{so} is denoted the articulation angle, and is defined as $\psi = \int r dt - \int r' dt$. The reference system S_{ts} describes a pitch and roll motion with respect to S_{to}, which is represented with the pitch angle χ_t and the roll angle ϕ_t. Correspondingly, S_{ss} describes a pitch and roll motion with respect to S_{so}, represented by χ_s and ϕ_s. The reference system S_h describes pitch and roll motions with respect to S_{to}, which are represented by χ_t and ϕ_h.

In the modeling of non-articulated vehicles it is common to express the equations of motion in coordinate systems attached to the sprung body [Kiencke and Nielsen, 2000]. For articulated vehicles, such as the tractor-semitrailer combination, this may lead to difficulties in reducing the equations using the kinematic articulation constraints. For the tractor-semitrailer model the equations of motion are most conveniently expressed in the S_{to} and S_{so} main reference systems aligned with the unsprung tractor and semitrailer frames, and fixed with respect to roll and pitch. This particular choice of reference system results in reasonable small expressions that may be handled by computer-algebra tools.

3. Kinematics

In the derivation of the equations of motion for the vehicle it is necessary to have expressions for the acceleration of arbitrary points on the vehicle body. Expressions for positions and velocities are needed for the suspension and tire models.

Coordinate transformations

Let $\bar{q}^{ts}$ be a vector representation in the tractor sprung body reference system S_{ts}. Then $\bar{q}^{to} = \underline{R}^{to}_{ts}\bar{q}^{ts}$ is the representation of this vector in main reference system S_{to}. The rotational transformation $\underline{R}^{to}_{ts}$ is defined as consecutive pitch and roll transformations (see Appendix B). Correspondingly, a vector $\bar{q}^{ss}$ in the semitrailer system S_{ss} is represented in S_{so} as $\bar{q}^{so} = \underline{R}^{so}_{ss}\bar{q}^{ss}$. Transformations between the semitrailer system S_{so} and the tractor system S_{to} are described with $\bar{q}^{to} = \underline{R}^{to}_{so}\bar{q}^{so}$.

The hitch kinematics is such that the roll angle ϕ_h of the tractor rear part $\mathcal{B}_h$ is given by

$$\phi_h = \phi_s \cos\psi + \chi_s \sin\psi \tag{1}$$

Thus it is equal to the roll angle of the semitrailer at zero articulation angle, and equal to the semitrailer pitch angle at 90 degrees articulation angle. The tractor rear part $\mathcal{B}_h$ always has the same pitch angle as the tractor front part $\mathcal{B}_{ts}$. Transformations between the tractor rear part system S_h and the tractor system S_{to} are described with $\bar{q}^{to} = \underline{R}^{to}_{h}\bar{q}^{h}$.

Motion of points on the vehicle bodies

Denote with S_* the earth-fixed inertial reference frame, and with S_i a vehicle fixed non-inertial reference frame rotating with the angular velocity $\bar{\omega}_i$ and translating with the velocity $\bar{v}_O$. Then for any vector $\bar{q}$ it holds that

$$\left(\frac{d}{dt}\right)_* \bar{q} = \left(\frac{d}{dt}\right)_i \bar{q} + \bar{\omega}_i \times \bar{q} \tag{2}$$

where '$\times$' denotes the cross-product operator. In particular the velocity of a point P with respect to S_* is expressed as the sum of the translational velocity of the vehicle reference system, and the time derivative of the position vector for P, $\bar{r}_P$:

$$\bar{v}_P = \bar{v}_O + \left(\frac{d}{dt}\right)_* \bar{r}_P = \bar{v}_O + \left(\frac{d}{dt}\right)_i \bar{r}_P + \bar{\omega}_i \times \bar{r}_P = \bar{v}_O + \dot{\bar{r}}_P + \bar{\omega}_i \times \bar{r}_P \tag{3}$$

The acceleration of P is computed by applying (2) on (3):

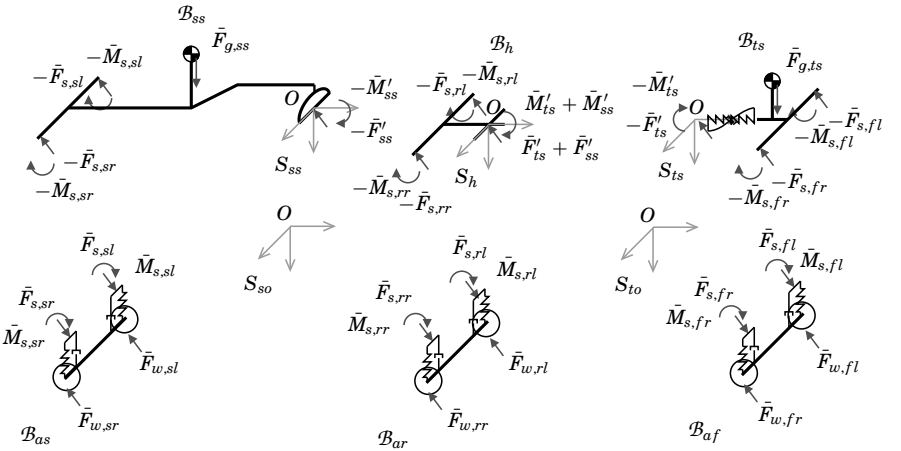

Figure 4 Vehicle free-body diagram.

$$\bar{a}_P = \left(\frac{d}{dt}\right)_* \bar{v}_P = \left(\frac{d}{dt}\right)_i (\bar{v}_O + \dot{\bar{r}}_P + \bar{\omega}_i \times \bar{r}_P) + \bar{\omega}_i \times (\bar{v}_O + \dot{\bar{r}}_P + \bar{\omega}_i \times \bar{r}_P)$$
$$= \dot{\bar{v}}_O + \bar{\omega}_i \times \bar{v}_O + \dot{\bar{\omega}}_i \times \bar{r}_P + \bar{\omega}_i \times (\bar{\omega}_i \times \bar{r}_P) + 2\bar{\omega}_i \times \dot{\bar{r}}_P + \ddot{\bar{r}}_P \quad (4)$$

Let $\bar{r}_P^{ts}$ denote the position of an arbitrary point P on the tractor body $\mathcal{B}_{ts}$, expressed in the tractor fixed reference system S_{ts}. Using the above relations, position $\bar{r}_P^{to}$, velocity $\bar{v}_P^{to}$ and acceleration $\bar{a}_P^{to}$ of P expressed in the tractor main reference system S_{to} are computed as in Appendix B. Expressions for the position, velocity and acceleration of points on the semitrailer body expressed in the semitrailer main system S_{so} are computed analogously. Likewise computed are the position and velocity of points on the hitch body $\mathcal{B}_h$.

4. Kinetics

The free-body diagram of Figure 4 introduces the forces and moments that are used in the kinematic analysis. The bodies $\mathcal{B}_{ts}$ and $\mathcal{B}_{ss}$ carry the masses m_{ts} and m_{ss}, with the inertial tensors $\underline{I}_{ts} = \int_{\mathcal{B}_{ts}} (r_P^2 \underline{1} - \bar{r}_P \bar{r}_P^T) dm_P$ and $\underline{I}_{ss} = \int_{\mathcal{B}_{ss}} (r_P^2 \underline{1} - \bar{r}_P \bar{r}_P^T) dm_P$ with respect to O.

The fundamental equations of dynamics in Euler's formulation postulates that

$$\bar{F} \equiv \int_{\mathcal{B}} d\bar{F}_P = \int_{\mathcal{B}} \bar{a}_P dm_P \quad (5a)$$

$$\bar{M}_O \equiv \int_{\mathcal{B}} \bar{r}_p \times d\bar{F}_P = \int_{\mathcal{B}} \bar{r}_P \times \bar{a}_P dm_P \tag{5b}$$

Thus for the tractor

$$\bar{F}_{ts} - \bar{F}'_{ts} = \int_{\mathcal{B}_{ts}} \bar{a}_P dm_P \tag{6a}$$

$$\bar{M}_{ts} - \bar{M}'_{ts} = \int_{\mathcal{B}_{ts}} \bar{r}_P \times \bar{a}_P dm_P \tag{6b}$$

where $\bar{F}_{ts}$ and $\bar{M}_{ts}$ are the sum of external moments and forces acting on $\mathcal{B}_{ts}$, and $\bar{F}'_{ts}$ and $\bar{M}'_{ts}$ are the internal reaction forces and moments from $\mathcal{B}_h$. Accordingly for the semitrailer

$$\bar{F}_{ss} - \bar{F}'_{ss} = \int_{\mathcal{B}_{ss}} \bar{a}_P dm_P \tag{7a}$$

$$\bar{M}_{ss} - \bar{M}'_{ss} = \int_{\mathcal{B}_{ss}} \bar{r}_P \times \bar{a}_P dm_P \tag{7b}$$

The static kinetic constraints that arise from the massless free body $\mathcal{B}_h$, are

$$\bar{F}'_{ts} + \bar{F}'_{ss} + \bar{F}_h = 0 \tag{8a}$$

$$\bar{M}'_{ts} + \bar{M}'_{ss} + \bar{M}_h = 0 \tag{8b}$$

where F_h and M_h are the external forces and moments acting on $\mathcal{B}_h$.

Elimination of internal forces and moments

The internal forces $\bar{F}'_{ts}$ and $\bar{F}'_{ss}$ are eliminated by combining (6a), (7a) and (8a), yielding

$$\bar{F}^{to}_{ts} + \underline{R}^{to}_{so} \bar{F}^{so}_{ss} + \bar{F}^{to}_h = \int_{\mathcal{B}_{ts}} \bar{a}^{to}_P dm_P + \underline{R}^{to}_{so} \int_{\mathcal{B}_{ss}} \bar{a}^{so}_P dm_P \tag{9}$$

when represented in S_{to}.

A little more care has to be taken when eliminating internal moments because of the torque transfer characteristics of the hitch. Regard that the hitch transfer moment to the semitrailer along the $\bar{e}^{ss}_x$ axis at 0 degrees articulation angle, and along the $\bar{e}^{ss}_y$ axis at 90 degrees articulation angle:

$$\bar{M}'^{ss}_{ss} = M'^{ss}_{ss,xy} \left(\cos\psi, \quad \sin\psi, \quad 0 \right)^T = M'^{ss}_{ss,xy} \cos\psi \bar{e}^{ss}_x + M'^{ss}_{ss,xy} \sin\psi \bar{e}^{ss}_y \tag{10}$$

The moment transfer to the tractor is in all $\bar{e}_x^{ts}$, $\bar{e}_y^{ts}$ and $\bar{e}_z^{ts}$ directions, and is determined by the roll stiffness in the $\bar{e}_x^{ts}$ direction:

$$\bar{M}'^{ts}_{ts} = \left(C_c \left(\phi_t - \phi_h \right), \quad M'^{ts}_{ts,y}, \quad M'^{ts}_{ts,z} \right)^T = C_c \left(\phi_t - \phi_h \right) \bar{e}_x^{ts} + M'^{ts}_{ts,y} \bar{e}_y^{ts} + M'^{ts}_{ts,z} \bar{e}_z^{ts} \quad (11)$$

Now (8b) combined with (10) and (11) may be expressed in S_{to} as

$$\underline{R}_{ts}^{to} \left(C_c \left(\phi_t - \phi_h \right) \bar{e}_x^{ts} + M'^{ts}_{ts,y} \bar{e}_y^{ts} + M'^{ts}_{ts,z} \bar{e}_z^{ts} \right) + \underline{R}_{ss}^{to} \left(M'^{ss}_{ss,xy} \cos\psi \bar{e}_x^{ss} + M'^{ss}_{ss,xy} \sin\psi \bar{e}_y^{ss} \right) + \underline{R}_h^{to} \bar{M}_h^h = 0 \quad (12)$$

or

$$\left(\underline{R}_{ss}^{to} \left(\cos\psi \bar{e}_x^{ss} + \sin\psi \bar{e}_y^{ss} \right) \quad \underline{R}_{ts}^{to} \bar{e}_y^{ss} \quad \underline{R}_{ss}^{to} \bar{e}_z^{ss} \right) \begin{pmatrix} M'^{ss}_{ss,xy} \\ M'^{ts}_{ts,y} \\ M'^{ts}_{ts,z} \end{pmatrix} = -\underline{R}_h^{to} \bar{M}_h^h - C_c \left(\phi_t - \phi_h \right) \underline{R}_{ts}^{to} \bar{e}_x^{ts} \quad (13)$$

which can be solved for $M'^{ss}_{ss,xy}$, $M'^{ts}_{ts,y}$ and $M'^{ts}_{ts,z}$. Thus (6b) and (7b) may be represented as

$$\bar{M}'^{to}_{ts} - \underline{R}_{ts}^{to} \bar{M}'^{ts}_{ts} = \int_{\mathcal{B}_{ts}} \bar{r}_P^{to} \times \bar{a}_P^{to} dm_P \quad (14)$$

$$\bar{M}'^{so}_{ss} - \underline{R}_{ss}^{so} \bar{M}'^{ss}_{ss} = \int_{\mathcal{B}_{ss}} \bar{r}_P^{so} \times \bar{a}_P^{so} dm_P \quad (15)$$

Equations (9), (14) and (15) form the complete set of equations of motion for the tractor-semitrailer combination vehicle.

State equations

Equations (9), (14) and (15) may be written in matrix form as

$$\begin{pmatrix} \bar{F}_1^{to} + \underline{R}_{so}^{to} \bar{F}_{ss}^{so} + \bar{F}_h^{to} \\ \bar{M}_{ts}^{to} - \underline{R}_{ts}^{to} \bar{M}'^{to}_{to} \\ \bar{M}_{ss}^{so} - \underline{R}_{ss}^{so} \bar{M}'^{ss}_{ss} \end{pmatrix} = \begin{pmatrix} \int_{\mathcal{B}_{ts}} \bar{a}_P^{to} dm_P + \underline{R}_{so}^{to} \int_{\mathcal{B}_{ss}} \bar{a}_P^{so} dm_P \\ \int_{\mathcal{B}_{ts}} \bar{r}_P^{to} \times \bar{a}_P^{to} dm_P \\ \int_{\mathcal{B}_{ss}} \bar{r}_P^{so} \times \bar{a}_P^{so} dm_P \end{pmatrix} \quad (16)$$

Introducing the state vector $\bar{\xi}^T = (\bar{\xi}_1^T, \bar{\xi}_2^T)$ with $\bar{\xi}_1^T = (U, V, W, r, \psi, \dot{\phi}_t, \dot{\chi}_t, \dot{\phi}_s, \dot{\chi}_s)$ and $\bar{\xi}_2^T = (Z, \psi, \phi_t, \chi_t, \phi_s, \chi_s)$, the right-hand side of (16) may be

rewritten such as

$$\begin{pmatrix} \int_{\mathcal{B}_{ts}} \bar{a}_P^{to} dm_P + \underline{R}_{so}^{to} \int_{\mathcal{B}_{ss}} \bar{a}_P^{so} dm_P \\ \int_{\mathcal{B}_{ts}} \bar{r}_P^{to} \times \bar{a}_P^{to} dm_P \\ \int_{\mathcal{B}_{ss}} \bar{r}_P^{so} \times \bar{a}_P^{so} dm_P \end{pmatrix} = \underline{H}(\bar{\xi}) \frac{d\bar{\xi}_1}{dt} - \underline{F}(\bar{\xi}) \tag{17}$$

where the elements of the matrices $\underline{H}(\bar{\xi})$ and $\underline{F}(\bar{\xi})$ are nonlinear functions of the state vector where inertial and geometrical parameters appear as constants. By combining (16) and (17) the complete state equations are

$$\underline{H}(\bar{\xi}) \frac{d\bar{\xi}_1}{dt} = \underline{F}(\bar{\xi}) + \begin{pmatrix} \bar{F}_{ts}^{to} + \underline{R}_{so}^{to} \bar{F}_{ss}^{so} + \bar{F}_h^{to} \\ \bar{M}_{ts}^{to} - \underline{R}_{ts}^{to} \bar{M}'^{ts}_{ts} \\ \bar{M}_{ss}^{so} - \underline{R}_{ss}^{so} \bar{M}'^{ss}_{ss} \end{pmatrix} \tag{18a}$$

$$\frac{d\bar{\xi}_2}{dt} = \underline{E} \bar{\xi}_1 \tag{18b}$$

with $\underline{E}$ defined in concordance with the definitions of $\bar{\xi}_1$ and $\bar{\xi}_2$. The matrices $\underline{H}(\bar{\xi})$ and $\underline{F}(\bar{\xi})$ are computed by inserting the kinematic expressions for positions and accelerations from Section 3 in the integrals, solving the integrals, and then factoring out the time derivatives of $\bar{\xi}_1$. The integral expressions are very large, but easy to solve with a computer-algebra tool as they evaluate to body masses, center of gravity locations and elements from the inertia matrices. The equations of motion are too large to be written out explicitly. The 9×9 matrix $\underline{H}(\bar{\xi})$ consists of elements of up to nine nonlinear terms in the state variables, and the 9×1 matrix $\underline{F}(\bar{\xi})$ of elements of up to 59 nonlinear terms in the state variables. Equation (18a) is solved numerically for the state derivatives at each integration step t_k, using the current state $\bar{\xi}(t_k)$. Solving for the state derivatives analytically is not practical, since it would lead to extremely complex expressions. For optimizing computational speed in simulations this could, however, be done using methods described in e.g. [Sayers, 1990]. The gain in simulation speed is still expected to be minor since solving the equation takes less than 0.5 ms on a standard Pentium II 366 MHz PC using LAPACK algorithms.

The external forces that act on the vehicle are the gravitational forces $\bar{F}_{g,ts} = m_{ts} g \bar{e}_z^{to}$ and $\bar{F}_{g,ss} = m_{ss} g \bar{e}_z^{so}$, and tire-road contact forces. Wind-drag forces are not included in the model, but may easily be added. The tire-road contact forces are obtained from the tire model, which is described below. These forces are transfered to the sprung chassis through the suspension, and depend on suspension characteristics and axle configurations. The suspension system determines the vertical forces as well as pitch and roll moments.

Axle balance equations

The axle and suspension kinetics are derived from analyzing the free-body diagram of Figure 5. The description of the axle-kinetics is presented for an arbitrary axle. The indices "l" and "r" on the variables denote the left or right side of axle. The axle-kinetics are described in the S_{to} and S_{so} reference systems for the tractor and semitrailer axles respectively. Denote the suspension forces at corner $i \in \{l, r\}$ by $\bar{F}_{s,i}$, the suspension moments by $\bar{M}_{s,i}$, and the wheel forces by $\bar{F}_{w,i}$. Now $F_{s,i,z}$, $F_{w,i,x}$, $F_{w,i,y}$, $M_{s,i,x}$, $M_{s,i,z}$ may be expressed directly in terms of position and velocity of the suspension corner, which can be computed from the state variables, while $F_{s,i,x}$, $F_{s,i,y}$, $F_{w,i,z}$, $M_{s,i,y}$ are computed by solving balance equations described below. Since the axle and the sprung chassis are rigid bodies that have two supporting connections aligned in the $\bar{e}_y^{to}$, $\bar{e}_y^{so}$ directions, there are no internal moments acting in the $\bar{e}_x^{to}$, $\bar{e}_x^{so}$ and $\bar{e}_z^{to}$, $\bar{e}_z^{so}$ directions. With a roll-stiffener an extra moment in the $\bar{e}_x^{to}$, $\bar{e}_x^{so}$ direction may be added. Let $\bar{M}_{s,x} = \bar{M}_{s,l,x} + \bar{M}_{s,r,x}$ and $\bar{M}_{s,y} = \bar{M}_{s,l,y} + \bar{M}_{s,r,y}$ denote the lumped suspension moments on one axle. This reduces the number of free variables. The planar tire-forces $F_{w,l,x}$, $F_{w,l,y}$, $F_{w,r,x}$, $F_{w,r,y}$ are assumed to depend linearly on the normal tire-forces $F_{w,l,z}$, $F_{w,r,z}$ as:

$$F_{w,l,x} = F_{w,l,z}\mu_{l,x} \qquad F_{w,l,y} = F_{w,l,z}\mu_{l,y} \tag{19a}$$

$$F_{w,r,x} = F_{w,r,z}\mu_{r,x} \qquad F_{w,r,y} = F_{w,r,z}\mu_{r,y} \tag{19b}$$

where the coefficients of friction $\mu_{l,x}$, $\mu_{l,y}$, $\mu_{r,x}$, and $\mu_{r,y}$ are given from the tire model. Let $\Delta r_{s,l,z}$, $\Delta r_{s,r,z}$, $v_{s,l,z}$, and $v_{s,r,z}$ denote the suspension travel and velocity at the left and right suspension corners at a given time. These variables are computed from the state variables using the expressions of Section 3. The vertical suspension forces may now be computed using any desired mapping. In this presentation a simple linear suspension is used:

$$F_{s,l,z} = -C_s\Delta r_{s,l,z} - D_s v_{s,l,z} \tag{20a}$$

$$F_{s,r,z} = -C_s\Delta r_{s,r,z} - D_s v_{s,r,z} \tag{20b}$$

The known suspension moments are $M_{s,x} = -C_r(\Delta r_{s,r,z} - \Delta r_{s,l,z})$ and $M_{s,z} = 0$, where an anti-roll bar with stiffness C_r provides additional roll stiffness.

Balance of forces and moments result in the following set of equations:

$$F_{s,l,x} + F_{s,r,x} + F_{w,l,x} + F_{w,r,x} = 0 \tag{21a}$$

$$F_{s,l,y} + F_{s,r,y} + F_{w,l,y} + F_{w,r,y} = 0 \tag{21b}$$

$$F_{s,l,z} + F_{s,r,z} + F_{w,l,z} + F_{w,r,z} = 0 \tag{21c}$$

$$-r_{s,l,z}F_{s,l,y} - r_{s,r,z}F_{s,r,y} + r_{s,l,y}F_{s,l,z} + r_{s,r,y}F_{s,r,z}$$

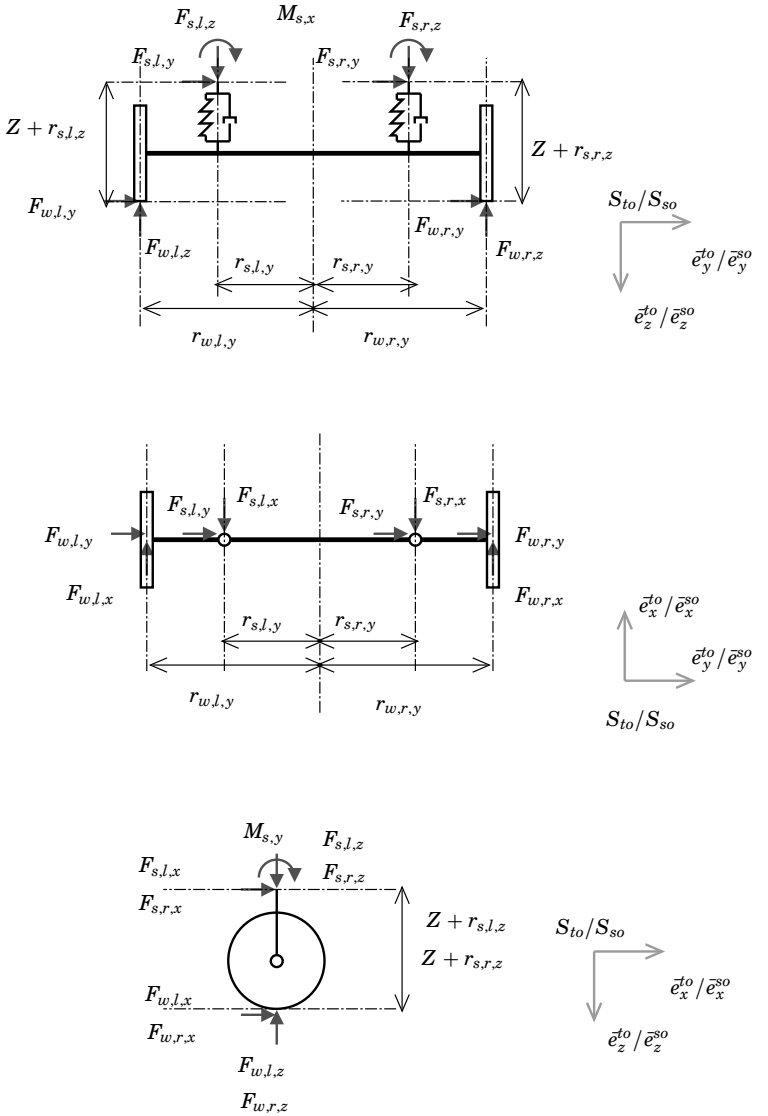

Figure 5 Tractor front axle and suspension kinetics. Tractor rear axle and semi-trailer axle have corresponding configurations.

$$+r_{w,l,y}F_{w,l,z} + r_{w,r,y}F_{w,r,z} + M_{s,x} = 0 \quad (21d)$$

$$r_{s,l,z}F_{s,l,x} + r_{s,r,z}F_{s,r,x} + M_{s,y} = 0 \quad (21e)$$

$$-r_{s,l,y}F_{s,l,x} - r_{s,r,y}F_{s,r,x} - r_{w,l,y}F_{w,l,x} - r_{w,r,y}F_{w,r,x} = 0 \quad (21f)$$

$$F_{s,l,y} = F_{s,r,y} \quad (21g)$$

It is assumed that side forces are equally carried by the left and right suspension (21g). Using the assumption (19) the Equations (21a–21g) can be rewritten as

$$\begin{pmatrix} 1 & 1 & 0 & 0 & \mu_{l,x} & \mu_{r,x} & 0 \\ 0 & 0 & 1 & 1 & \mu_{l,y} & \mu_{r,y} & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -r_{s,l,z} & -r_{s,r,z} & r_{w,l,y} & r_{w,r,y} & 0 \\ r_{s,l,z} & r_{s,r,z} & 0 & 0 & 0 & 0 & 1 \\ r_{s,l,y} & r_{s,r,y} & 0 & 0 & r_{w,l,y}\mu_{l,x} & r_{w,r,y}\mu_{r,x} & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} F_{s,l,x} \\ F_{s,r,x} \\ F_{s,l,y} \\ F_{s,r,y} \\ F_{w,l,z} \\ F_{w,r,z} \\ M_{s,y} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -F_{s,l,z} - F_{s,r,z} \\ -M_{s,x} - r_{s,l,y}F_{s,l,z} - r_{s,r,y}F_{s,r,z} \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (22)$$

This linear equation system is solved for each axle at each simulation step.

During heavy cornering the right or left wheel may not stay in contact with the ground. The normal force of the left or right wheel is then zero, i.e. $F_{w,l,z} = 0$ or $F_{w,r,z} = 0$, and is no longer a free variable in Equation (22). The axle roll angle may then be introduced as a new free variable. A new balance equation may then be stated and solved [Gäfvert and Lindgärde, 2001].

For certain tires the friction coefficients in (19) decrease at large loads, so that the planar tire-forces depend nonlinearly on the normal forces. If the friction coefficients depend on the normal forces the difficulty arises that tire normal forces need to be known to compute tire lateral and longitudinal forces, while the lateral and longitudinal forces need to be known to solve the axle balance equations for the normal forces. In this case it is still possible to use the axle-balance equations 22 iteratively.

Then the equations are first solved using a nominal load, and the resulting vertical forces are used in the next iteration. The convergence is in general very fast. In tested cases two iterations have been sufficient.

5. Tire and Wheel Models

The model presented below is based on the generic *slip circle* model [Schuring *et al.*, 1996, Pottinger *et al.*, 1998] for combined braking and cornering. The basic idea is to combine models for pure braking and pure cornering into a model that accurately describes simultaneous braking and cornering. The tire model describes the longitudinal and lateral tire-road contact forces. Effects with limited influence on vehicle handling are neglected. This includes self-aligning torque which mainly influences steering dynamics [Edlund, 1991], effects of camber angle which is normally small on trucks [Pauwelussen *et al.*, 1999], and rolling resistance.

The model is expressed in terms of tire slips. Longitudinal tire slip λ is defined as

$$\lambda = \frac{\omega R_w - u}{u} \tag{23}$$

where u, v are the longitudinal and lateral velocities of the wheel center expressed in a reference system aligned with the wheel, ω is the wheel angular velocity, and R_w is the effective wheel radius. The lateral side-slip α is defined as

$$\tan\alpha = -\frac{v}{u} \tag{24}$$

A dimensionless slip variable s is introduced as

$$s = \sqrt{\lambda^2 + \sin^2\alpha} \tag{25}$$

Define the slip angle β as

$$\tan\beta = \frac{\sin\alpha}{\lambda} \tag{26}$$

The tire force is assumed to be counter-directed to the slip vector described by s and β. The longitudinal and lateral adhesion coefficients μ_x and μ_y are defined as

$$\mu_x = \frac{F_{w,x}}{F_{w,z}} \tag{27a}$$

$$\mu_y = \frac{F_{w,y}}{F_{w,z}} \tag{27b}$$

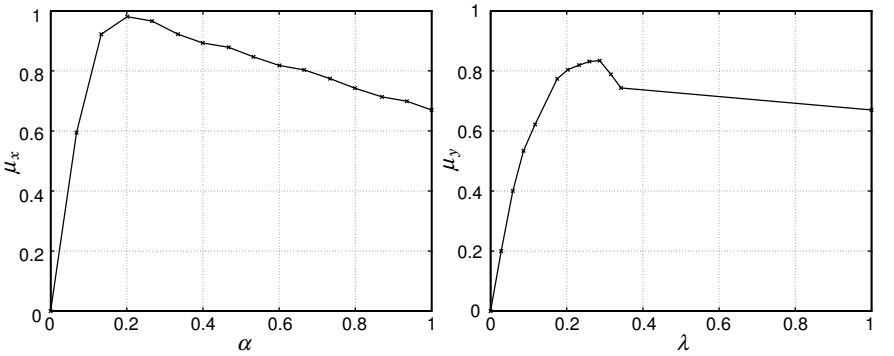

Figure 6 Experimental tire adhesion coefficients at pure braking (left) and pure cornering (right). The last data value for the lateral adhesion coefficient is estimated as being equal to the longitudinal adhesion coefficient for $\lambda = 1$. The experimental data is obtained from Volvo Truck Corporation [Edlund, 1991].

where $F_{w,z}$ is the tire vertical load, and where $F_{w,x}$ and $F_{w,y}$ are the longitudinal and lateral tire forces. The adhesion coefficients for pure cornering and braking are modeled as mappings from the tire slips as $\mu_x(\lambda)$ and $\mu_y(\alpha)$. The adhesion coefficients may also be dependent on the vertical load. Experimentally acquired tire data is used for the maps (see Figure 6). The adhesion coefficient of the combined tire force is now described as

$$\mu(s,\beta) = \mu_x(s)\cos^2\beta + \mu_y(s)\sin^2\beta \tag{28}$$

Finally, the resulting tire forces are obtained from

$$F_{w,x} = F_{w,z}\mu(s,\beta)\cos\beta \tag{29a}$$

$$F_{w,y} = F_{w,z}\mu(s,\beta)\sin\beta \tag{29b}$$

It is seen that pure cornering and braking are restored for $\beta = 0$ and $\beta = \pi/2$. For other β the combined tire forces lie on a curve that is "close" to an ellipse. The slip circle model is compared with experiments in Figure 7. At combined braking and cornering the adhesion coefficients are reduced in contrast to pure braking and pure cornering.

The rotational dynamics of the wheels are modeled as

$$J_w\frac{d}{dt}\omega = M_\omega - R_w F_{w,x} \tag{30}$$

where J_w is the moment of inertia of the wheel, and M_w the driving or braking moment. For certain applications where brake and wheel dynamics are not of particular interest, the wheel dynamics may be eliminated

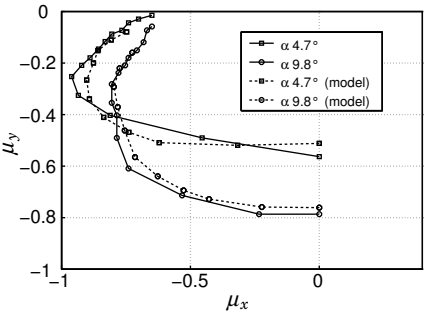

Figure 7 Slip circle model and experimental adhesion coefficients at combined braking and cornering for two fixed α and varying λ. The experimental data is obtained from Volvo Truck Corporation [Edlund, 1991].

and the longitudinal tire slips λ may then be used directly as inputs to the vehicle model.

A simple model of steering compliance may be added for each steered wheel as $\delta = K_{\text{steer}}\delta_{\text{steer}} + M_z/C_{\text{steer}}$, where δ is the steering angle of the wheel, K_{steer} the geometric steering ratio, δ_{steer} the steering-wheel angle, C_{steer} the steering compliance, and M_z the self-aligning torque acting on the wheel. For details and more elaborate models, see e.g. [Gillespie, 1992]. With this amendment one may study responses to steering-wheel inputs instead of direct wheel angle inputs. It is then necessary to use a tire model which also describes the self-aligning torque, e.g. [Bakker *et al.*, 1989].

Roll-steer effects have significant influence on handling performance [Sweatman and McFarlane, 2000], and is easily included as describing wheel steering angles as functions of the roll angle, e.g. $\delta = K_{\text{roll}}\phi_h$, where K_{roll} is the roll-steer coefficient.

6. Implementation

A model of a 4x2 tractor with a 3-axle semitrailer was implemented in the Matlab-Simulink simulation environment. The computer algebra software Maple was used to compute $\underline{F}(\bar{\xi})$, $\underline{H}(\bar{\xi})$ and the kinematic relations of Section 3, see Appendix C. Maple has support for translation of expressions to the programming language C. The expressions for $\underline{F}(\bar{\xi})$ resulted in around 300 lines of C-code, and the expressions for $\underline{H}(\bar{\xi})$ in around 200 lines. The generated C-code was included in Matlab mex-file functions which are called during simulations. The resulting simulation code

is modular in the sense that the dynamic equations, the suspensions and the axle load balance equations, and the tire model equations are separated. This makes it easy to substitute one of these sub-models for another. Wheel rotational dynamics are not included. Hence, the longitudinal slips λ_i with $i \in \{fl, fr, rl, rr, s1l, s1r, s2l, s2r, s3l, s3r\}$, and the front wheel steering angle δ are the inputs to the model. The parameters of the model are listed in Appendix D. A 3D-animation routine based on standard Matlab graphics was developed to illustrate simulation results. This provides good supplement to the signal graphs and helps understanding the vehicle behaviour in different scenarios.

7. Results

Validation

Experimental data recorded from 31 runs on a test-track with a 4x2 tractor-semitrailer vehicle was provided by Volvo Truck Corporation, together with data of the vehicle. A set of model parameters corresponding to the tested vehicle was derived. Since the provided data did not cover the complete set of parameters of the model, some assumptions were made. In particular it was difficult to determine tire data and suspension data with high accuracy. No information except from approximate mass and dimensions was provided on the semitrailer parameters. Nonlinear lookup-table mappings for the damper characteristics were provided, and were included in the model. Braking actions were not recorded, and most of the test-scenarios only included steering actions. The recorded variables of interest were: front wheel angles, vehicle speed, kingpin-angle, yaw-rate, and suspension travel for the four corners of the tractor.

A linear constant speed planar 3-DOF dynamic tractor-semitrailer model [Gäfvert, 2001] was tuned with parameters corresponding to those of the 9-DOF model. This 3-DOF model has the lateral velocity V, yaw-rate r, articulation angular rate $\dot{\psi}$, and articulation angle ψ as state variables. Both the 3-DOF and the 9-DOF models were simulated with the recorded front steering angle of the experimental data as inputs. Figure 8 shows the typical results from a lane-change maneuver. Both models show good accordance with experimental data with respect to kingpin-angle and yaw-rate. The suspension travel of the 9-DOF model is well reproduced. The tire parameters have large impact on the reproduction of the king-pin angle and the yaw-rate, and need careful tuning.

The motivation for using the 9-DOF model instead of the simpler 3-DOF model is the inclusion of the load-transfer effects on tire normal

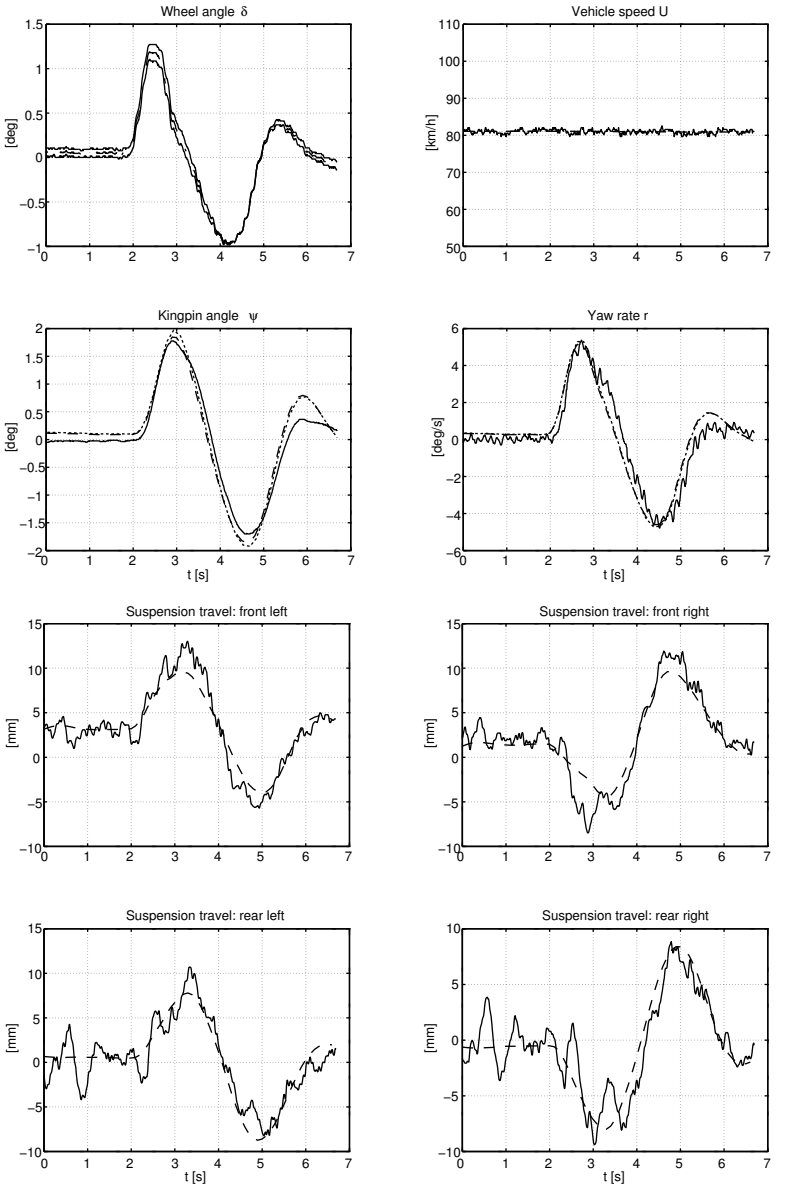

Figure 8 Solid: Experiment; Dashed: 9-DOF model; Dotted: Linear 3-DOF model. In the upper four plots the dashed and dotted lines almost coincide. Note that the linear 3-DOF model does not reproduce the suspension travel of the four lower diagrams in the figure.

forces. Those forces were not measured in the experiments, but they correspond to the suspension travel. Since the suspension travel is reasonably well reproduced by the 9-DOF model it is concluded that the tire normal forces are equally well reproduced. In the present scenarios the linear model and the 9-DOF model perform similarly with respect to kingpin-angle and yaw-rate. Hence, good reproduction of normal forces are not needed to reproduce the planar behavior of these scenarios. This may be explained by the fact that in pure cornering only the lateral tire-force produced at each axle is of importance, i.e. the sum of the left and right wheel lateral forces. The load transfer affects this total lateral force only in presence of significant nonlinear normal-force dependence in the tire models, having the effect of decreasing the lateral force at large loads. Still, the load transfer will have significant influence in the case of unilateral braking actions, which may be applied by yaw-stabilization systems. This is because the largest available planar tire-forces at each wheel are determined by the normal forces.

Simulation with stabilizing controller using unilateral braking

A linear 3-DOF constant-speed model of the tractor [Gäfvert, 2001] was used to design a simple yaw stabilizing controller using unilateral braking actions. The controller consists of two discrete-time LQ-type controllers [Åström and Wittenmark, 1997] that handle oversteer and understeer situations by braking a front axle or rear axle wheel on the tractor, correspondingly. The controllers minimize cost functions of the type

$$J = \int (\bar{x} - \bar{x}_r)^T \underline{Q} (\bar{x} - \bar{x}_r) + R\lambda^2 \, dt \tag{31}$$

where $\bar{x} = (V, r)$ is the lateral velocity and yaw rate of the tractor, $\bar{x}_r$ is the desired lateral velocity and yaw rate of the tractor computed from a linear reference model, λ is the controlled braking action described as longitudinal tire slip, and where $\underline{Q}$ and R are tuning parameters. Hence, the controller minimizes a trade-off, determined by the tuning parameters, between the integral of the squared control error in lateral speed and yaw, and the power of the braking interventions. Note that the ratio V/U is the vehicle side-slip β, for small values. Hence, controlling V is essentially the same as controlling the side-slip β. The resulting controller is of the form

$$\lambda_f = \underline{L}_{\text{oversteer}} (\bar{x}_r - \bar{x}) \tag{32a}$$

$$\lambda_r = \underline{L}_{\text{understeer}} (\bar{x}_r - \bar{x}) \tag{32b}$$

where $\underline{L}_{\text{oversteer}}$ and $\underline{L}_{\text{understeer}}$ are constant 1×2 matrices. When oversteer or understeer behaviour is detected the corresponding controller inter-

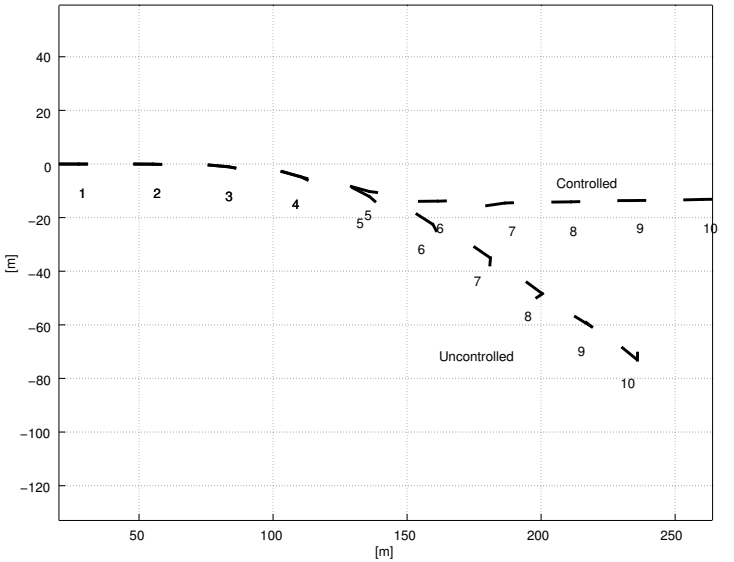

Figure 9 Paths for the uncontrolled and controlled vehicle, with the tractor and semitrailer illustrated as connected bars for $t = 1$ to 10 s. The maneuver results in a jackknifing situation for the uncontrolled vehicle.

venes with braking actions on the front or rear axle. The sign of λ determines wether the braking action is distributed to the left or right wheel. The sampling time of the controller is 50 ms. A similar control strategy is described in [van Zanten *et al.*, 1995, Hecker *et al.*, 1997].

The controller was applied to the 9-DOF model and simulated with a violent lane-change maneuver at 100 km/h and reduced road adhesion. In Figure 9 the effect of the controller is illustrated by showing the vehicle paths for the controlled and uncontrolled vehicle. Without the stabilizing controller the vehicle enters a jackknifing situation, while the controlled vehicle maintains stability. Figure 10 shows some primary vehicle variables during the maneuver. The front-wheel angle input is constructed from cosine quarter-periods to obtain a smooth trajectory. The lateral acceleration and the normal tire-forces show that the maneuver is rather violent, not very far from roll-over. As can be seen, the yaw-rate, vehicle side-slip, and articulation angle are quite rapidly stabilized after the lane-change. Figure 11 shows the braking actions commanded by the controller. Braking on the front left and front right wheels compensate the oversteering. In the transition from right to left turn a short braking action is also commanded on the rear left wheel. A snapshot of the animation of the controlled maneuver is shown in Figure 12. The simulation indicates the

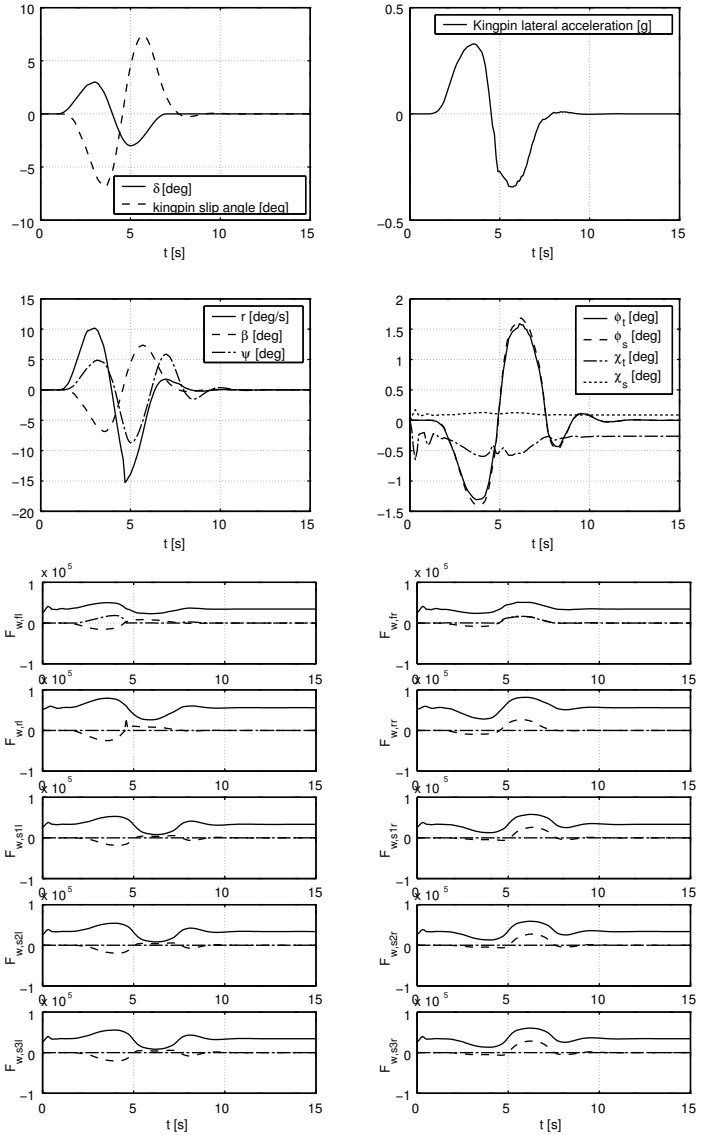

Figure 10 Controlled vehicle variables. Top left: front wheels angle, tractor slip angle; Top right: lateral acceleration; Middle left: yaw rate, articulation angle, vehicle side-slip; Middle right: tractor and semitrailer roll and pitch angles; Bottom: longitudinal (dash-dotted), lateral (dashed), and vertical (solid) tire forces.

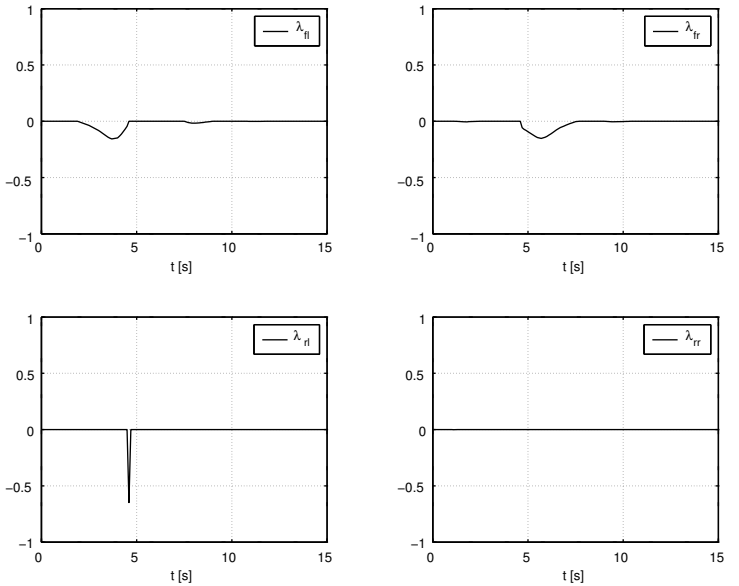

Figure 11 Longitudinal tire slips resulting from unilateral braking actions of the stabilizing controller.

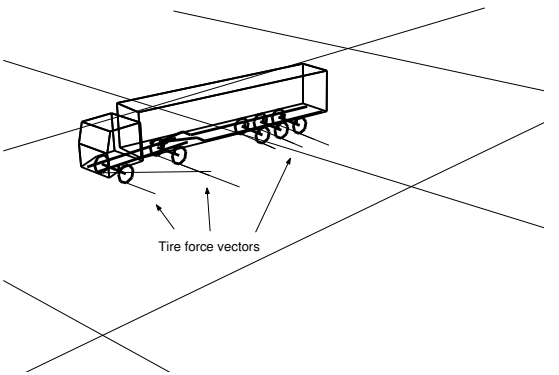

Figure 12 Snapshot of the animation of the controlled lane-change maneuver at time $t = 5.4$ s. Note the front right braking action in the tire force vectors.

applicability of the 9-DOF model to handling simulations for evaluating stabilizing control strategies. The controller used in this simulation is to be regarded as a simple example. More sophisticated controller designs need to be used to handle, e.g., the uncertainties in mass distribution and tire-road adhesion. Moreover, the signals used for feedback in the example are normally not directly available, but must be reconstructed by an observer.

8. Conclusions

The nonlinear dynamic 9-DOF tractor-semitrailer model presented in this paper is modular in the sense that axle, suspension, and tire sub-models are separate from the chassis dynamics. Hence, it is easy to change axle and wheel configurations to plug in different tire models, or to use different suspension models. Moreover, the model should be easy to extend with e.g. brake dynamics. The equations of motion are derived without other approximations than the assumption on the vehicle configuration. Approximations are introduced in the modeling of the suspensions, axles, and tires. The level of approximation may be chosen by the user by substituting the subsystem in question for a more detailed model. This is a simple operation due to the modularity.

Model parameters are intuitive and quite easy to tune. Geometry and mass parameters are easy to measure on the real vehicle. Mass distributions are usually known at least approximately. The model is robust to small variations in these parameters in the sense that the corresponding variations in simulation results are small. Suspension characteristics are normally known well. They primarily determine the load-transfer dynamics. In presence of highly nonlinear normal-load dependence of the tires they may have significant influence on the directional response in cornering. Tire characteristics are a difficult matter, not the least for combined braking and cornering, and have great influence on the directional response of the vehicle. Therefore, tire parameters need careful tuning. Sets of tire data from test-bench experiments at pure braking or cornering are often available, and may be used in combination with a suitable tire model. Even though all data were not available for the test vehicle in this study, good results were obtained by using estimates of unknown entities.

The validation indicates that load-transfer is accurately modeled. This implies that the model may give realistic results in simulations of handling maneuvers near and beyond the tire adhesion limits. Hence, the model is suitable for studies on advanced chassis control in handling-maneuvers, as shown in the example with the yaw stabilization controller.

A computer-algebra tool has been indispensable in the construction and implementation of the model, since the expressions are far too large to manipulate by hand with reasonable effort. As seen in appendix C, only a modest set of Maple commands are needed to derive the equations of motion. Hence, with a powerful computer-algebra tool and basic knowledge in analytical mechanics it is possible to derive a fairly complex model of the tractor-semitrailer vehicle without using specialized domain-specific modelling software.

Acknowledgements

This work is part of the DICOSMOS (Distributed Control of Safety-Critical Motion Systems) project, funded by the Swedish research board NUTEK under the Complex Technical Systems program, with project nr. P11762-2, and by Volvo Technological Development (VTD). Mats Andersson (VTD) was the initiator of the work on the semi-trailer combination vehicle. Niklas Fröjd at Volvo Truck Corporation has been very helpful with providing validation data, and general comments on modeling issues.

References

Anderson, D., G. Shade, S. Hamill, and P. O'Heron (2001): "Development of a multi-body dynamic model of a tractor-semitrailer for ride quality prediction." SAE Technical Paper 2001-01-2764.

Åström, K. J. and B. Wittenmark (1997): *Computer-Controlled Systems: Theory and Design*, 3 edition. Prentice Hall International.

Bakker, E., H. B. Pacejka, and L. Lidner (1989): "A new tire model with an application in vehicle dynamics studies." SAE Technical Paper 890007.

Chen, C. and M. Tomizuka (1995): "Dynamics modeling of articulated vehicles for automated highway systems." In *Proc. of the American Control Conference*, pp. 653–657. Seattle, Washington.

Chen, C. and M. Tomizuka (2000): "Lateral control of commercial heavy vehicles." *Vehicle System Dynamics*, **No 33**, pp. 391–420.

Edlund, S. (1991): "Tyre models: Subreport –91." Technical Report. Volvo Truck Corporation. (classified).

Ellis, J. (1994): *Vehicle Handling Dynamics*. Mechanical Engineering Publications Limited, London.

Ellis, J. R. (1969): *Vehicle Dynamics*. Business Books Ltd., London.

Ellis, J. R. (1988): *Road Vehicle Dynamics*. John R. Ellis Inc.

Fowles, G. R. and G. L. Cassiday (1993): *Analytical Mechanics*, 5 edition. Saunders College Publishing.

Gäfvert, M. (2001): "Studies on yaw-control of heavy-duty trucks using unilateral braking." Technical Report ISRN LUTFD2/TFRT–7598–SE. Department of Automatic Control, Lund Institute of Technology, Sweden. DICOSMOS Report.

Gäfvert, M. and O. Lindgärde (2001): "A 9-DOF tractor-semitrailer dynamic handling model for advanced chassis control studies." Technical Report ISRN LUTFD2/TFRT–7597–SE. Department of Automatic Control, Lund Institute of Technology, Sweden. DICOSMOS Report.

Gillespie, T. D. (1992): *Fundamentals of Vehicle Dynamics*. Society of Automotive Engineers.

Hecker, F., S. Hummel, O. Jundt, K.-D. Leimback, I. Faye, and H. Schramm (1997): "Vehicle dynamics control for commercial vehicles." SAE Technical Paper 973284.

Kiencke, U. and L. Nielsen (2000): *Automotive Control Systems*. Springer.

Leucht, P. M. (1970): "The directional dynamics of the commercial tractor-semitrailer vehicle during braking." SAE Technical Paper 700371.

Lidström, P. and U. Uhlhorn (1984): "Teknisk mekanik 2f.". Avdelningen för Mekanik, Lunds Tekniska Högskola. (In Swedish).

Mikulcik, E. C. (1968): *The dynamics of Tractor-Semitrailer Vehicles: The Jackknifing Problem*. PhD thesis, Cornell University.

Mikulcik, E. C. (1971): "The dynamics of tractor-semitrailer vehicles: The jackknifing problem." SAE Technical Paper 710045.

Pacejka, H. B. (1988): "Modelling of the pneumatic tyre and its impact on vehicle dynamic behavior." Technical Report i72B. Technische Universiteit Delft.

Pauwelussen, J., G. Anghelache, Theodoresco, and A. Schmeitz (1999): "Truck tyre behavior in use and testing methods." In *European Tyre School: Teaching Modules*. Tampere University of Technology.

Pottinger, M. G., W. Pelz, and G. Falciola (1998): "Effectiveness of the slip circle, "combinator", model for combined tire cornering and braking forces when applied to a range of tires." SAE Technical Paper 982747.

Ranganathan, R. and A. Aia (1995): "Development of heavy vehicle dynamic stability analysis model using MATLAB/SIMULINK." SAE Technical Paper 952638.

Sayers, M. (1990): *Symbolic Computer Methods to Automatically Formulate Vehicle Simulation Codes*. PhD thesis, The University of Michigan.

Sayers, M. W. (1992): "Computer modeling and simulation of vehicle dynamics." *UMTRI Research Review*, **23:1**, pp. 1–15.

Schuring, D. J., W. Pelz, and M. G. Pottinger (1996): "A model for combined tire cornering and braking forces." SAE Technical Paper 960180.

Sweatman, P. F. and S. McFarlane (2000): "Investigation into the specification of heavy trucks and consequent effects on truck dynamics and drivers: Final report." Technical Report. Department of Transport and Regional Services, Commonwealth of Australia.

van Zanten, A. T., R. Erhart, and G. Pfaff (1995): "VDC, the vehicle dynamics control system of Bosch." SAE Technical Paper 950759.

A. Nomenclature

Body configuration and reference systems	
$\mathcal{B}_{af}$	Tractor front axle unsprung body
$\mathcal{B}_{ar}$	Tractor rear axle unsprung body
$\mathcal{B}_{as}$	Semitrailer axle(s) unsprung body
$\mathcal{B}_{ts}$, S_{ts}	Tractor sprung body with attached reference system
$\mathcal{B}_{ss}$,S_{ss}	Semitrailer sprung body with attached reference system
$\mathcal{B}_{h}$, S_{h}	Hitch (tractor rear part) body with attached reference system
S_{to}	Tractor main reference system
S_{so}	Semitrailer main reference system
O	Origin for S_{to}, S_{so}, S_{ts}, S_{ss}, S_{h}

Indices	
$_P$	A point on the vehicle body
$_{to}$	Tractor main
$_{so}$	Semitrailer main
$_{af}$	Tractor unsprung front axle
$_{ar}$	Tractor unsprung rear axle
$_{as}$	Semitrailer unsprung axle(s)
$_{ts}$	Tractor sprung
$_{ss}$	Semitrailer sprung
$_{h}$	Hitch
$_{f}$	Front axle
$_{r}$	Rear axle
$_{s}$	Semitrailer axle
$_{l}$	Left side
$_{r}$	Right side

Geometry	
$\bar{r}_{CM,ts}$	Center of mass location for $\mathcal{B}_{ts}$
$\bar{r}_{CM,ss}$	Center of mass location for $\mathcal{B}_{ss}$
$\bar{r}_{s,i,j}$	Location of suspension at side j on axle i
$\bar{r}_{w,i,j}$	Location of wheel at side j on axle i

Kinematics	
$\bar{v}_O$	velocity of O expressed in S_{to}
U	Longitudinal component of v_O
V	Lateral component of v_O
W	Vertical component of v_O
$\bar{\omega}_{to}$	Angular velocity of S_{to}
r	Yaw velocity of S_{to}
$\bar{\omega}_{so}$	Angular velocity of S_{so}
r'	Yaw velocity of S_{so}
ψ	Articulation angle (angle between S_{to} and S_{so})
ϕ_t, ϕ_s	Tractor and semitrailer roll angles
χ_t, χ_s	Tractor and semitrailer pitch angles
ϕ_h	Hitch body roll angle
$\underline{R}_i^j$	Coordinate transformation matrix from system i to system j
λ	Longitudinal tire-slip
α	Tire slip-angle
β	Vehicle side-slip angle
δ	Wheel steering-angle
δ_{steer}	Steering-wheel angle

Kinetics	
m_{ts}	Tractor mass ($\mathcal{B}_{ts}$)
m_{ss}	Semitrailer mass ($\mathcal{B}_{ss}$)
$\underline{I}_{ts}$	Tractor inertial tensor with respect to O
$\underline{I}_{ss}$	Semitrailer inertial tensor with respect to O
$\bar{F}'_{ts}$	Internal force between $\mathcal{B}_{ts}$ and $\mathcal{B}_h$
$\bar{M}'_{ts}$	Internal moment between $\mathcal{B}_{ts}$ and $\mathcal{B}_h$
$\bar{F}_{ss}$	Internal force between $\mathcal{B}_{ss}$ and $\mathcal{B}_h$
$\bar{M}'_{ss}$	Internal moment between $\mathcal{B}_{ss}$ and $\mathcal{B}_h$
$\bar{F}_{ts}$	Sum of external forces acting on $\mathcal{B}_{ts}$
$\bar{M}_{ts}$	Sum of external moments acting on $\mathcal{B}_{ts}$
$\bar{F}_{ss}$	Sum of external forces acting on $\mathcal{B}_{ss}$
$\bar{M}_{ss}$	Sum of external moments acting on $\mathcal{B}_{ss}$
$\bar{F}_{s,i,j}$	Suspension force at side j on axle i.
$\bar{F}_{w,i,j}$	Tire-road contact force at wheel on side j of axle i
$M_{z,i,j}$	Tire self-aligning torque at wheel on side j of axle i

B. Kinematic Expressions

Transformation matrices

$$\underline{R}_{ts}^{to} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_t & -\sin\phi_t \\ 0 & \sin\phi_t & \cos\phi_t \end{pmatrix} \begin{pmatrix} \cos\chi_t & 0 & \sin\chi_t \\ 0 & 1 & 0 \\ -\sin\chi_t & 0 & \cos\chi_t \end{pmatrix}$$

$$\underline{R}_{ss}^{so} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_s & -\sin\phi_s \\ 0 & \sin\phi_s & \cos\phi_s \end{pmatrix} \begin{pmatrix} \cos\chi_s & 0 & \sin\chi_s \\ 0 & 1 & 0 \\ -\sin\chi_s & 0 & \cos\chi_s \end{pmatrix}$$

$$\underline{R}_{so}^{to} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\underline{R}_{h}^{to} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_h & -\sin\phi_h \\ 0 & \sin\phi_h & \cos\phi_h \end{pmatrix} \begin{pmatrix} \cos\chi_t & 0 & \sin\chi_t \\ 0 & 1 & 0 \\ -\sin\chi_t & 0 & \cos\chi_t \end{pmatrix}$$

Motion of a point on the tractor sprung body $\mathcal{B}_{ts}$

$$\begin{aligned} \bar{r}_P^{to} &= \underline{R}_{ts}^{to}\bar{r}_P^{ts} \\ \bar{v}_P^{to} &= \bar{v}_O^{to} + \underline{\dot{R}}_{ts}^{to}\bar{r}_P^{ts} + \bar{\omega}_{to}^{to} \times \underline{R}_{ts}^{to}\bar{r}_P^{ts} \\ \bar{a}_P^{to} &= \dot{\bar{v}}_O^{to} + \bar{\omega}_{to}^{to} \times \bar{v}_O^{to} + \underline{\ddot{R}}_{ts}^{to}\bar{r}_P^{ts} + \dot{\bar{\omega}}_{to}^{to} \times \underline{R}_{ts}^{to}\bar{r}_P^{ts} \\ &\quad + 2\bar{\omega}_{to}^{to} \times \underline{\dot{R}}_{ts}^{to}\bar{r}_P^{ts} + \bar{\omega}_{to}^{to} \times \left(\bar{\omega}_{to}^{to} \times \underline{R}_{ts}^{to}\bar{r}_P^{ts}\right) \end{aligned}$$

Motion of a point on the semitrailer sprung body $\mathcal{B}_{ss}$

$$\begin{aligned} \bar{r}_P^{so} &= \underline{R}_{ss}^{so}\bar{r}_P^{ss} \\ \bar{v}_P^{so} &= \bar{v}_O^{so} + \underline{\dot{R}}_{ss}^{so}\bar{r}_P^{ss} + \bar{\omega}_{so}^{so} \times \underline{R}_{ss}^{so}\bar{r}_P^{ss} \\ \bar{a}_P^{so} &= \dot{\bar{v}}_O^{so} + \bar{\omega}_{so}^{so} \times \bar{v}_O^{so} + \underline{\ddot{R}}_{ss}^{so}\bar{r}_P^{ss} + \dot{\bar{\omega}}_{so}^{so} \times \underline{R}_{ss}^{so}\bar{r}_P^{ss} + 2\bar{\omega}_{so}^{so} \times \underline{\dot{R}}_{ss}^{so}\bar{r}_P^{ss} \\ &\quad + \bar{\omega}_{so}^{so} \times (\bar{\omega}_{so}^{so} \times \underline{R}_{ss}^{so}\bar{r}_P^{ss}) \end{aligned}$$

Motion of a point on the hitch body $\mathcal{B}_h$

$$\begin{aligned} \bar{r}_P^{to} &= \underline{R}_{h}^{to}\bar{r}_P^{h} \\ \bar{v}_P^{to} &= \bar{v}_O^{to} + \underline{\dot{R}}_{h}^{to}\bar{r}_P^{h} + \bar{\omega}_{to}^{to} \times \underline{R}_{h}^{to}\bar{r}_P^{h} \end{aligned}$$

C. Maple code

```
> restart;
```

```
> with(LinearAlgebra):
============ KINEMATICS ============
Motion of a point in vehicle coordinates
> reldiff := (r_::Vector,omega_::Vector) -> map(s_->diff(s_,t),r_) +
  CrossProduct(omega_,r_);
> rp := <x,y,z>;
> `v_O^to` := <U(t),V(t),W(t)>;
> `omega_to^to` := <0,0,r(t)>;
> `omega_so^so` := <0,0,r(t)-diff(psi(t),t)>;
Coordinate transformations
> `R_ts^to` := Matrix([[1,0,0],
    [0,cos(`phi_t`(t)),-sin(`phi_t`(t))],
    [0,sin(`phi_t`(t)),cos(`phi_t`(t))]]).
  Matrix([[cos(`chi_t`(t)),0,sin(`chi_t`(t))],
    [0,1,0],
    [-sin(`chi_t`(t)),0,cos(`chi_t`(t))]]);
> `R_ss^so` := Matrix([[1,0,0],
    [0,cos(`phi_s`(t)),-sin(`phi_s`(t))],
    [0,sin(`phi_s`(t)),cos(`phi_s`(t))]]).
  Matrix([[cos(`chi_s`(t)),0,sin(`chi_s`(t))],
    [0,1,0],
    [-sin(`chi_s`(t)),0,cos(`chi_s`(t))]]);
> `R_so^to` := Matrix([[cos(psi(t)),sin(psi(t)),0],
  [-sin(psi(t)),cos(psi(t)),0],
  [0,0,1]]);
> `R_to^so` := map(simplify,MatrixInverse(`R_so^to`));
> `phi_h` := `phi_s`(t)*cos(psi(t))+`chi_s`(t)*sin(psi(t));
> `R_h^to` := Matrix([[1,0,0],
    [0,cos(`phi_h`),-sin(`phi_h`)],
    [0,sin(`phi_h`),cos(`phi_h`)]]).
  Matrix([[cos(`chi_t`(t)),0,sin(`chi_t`(t))],
    [0,1,0],
    [-sin(`chi_t`(t)),0,cos(`chi_t`(t))]]);
Motion of a point on the tractor body
> `r_Pts^ts` := <x,y,z>;
> `r_Pts^to` := `R_ts^to`.`r_Pts^ts`;
> `v_Pts^to` := `v_O^to` + reldiff(`r_Pts^to`,`omega_to^to`);
> `a_Pts^to` := reldiff(`v_Pts^to`,`omega_to^to`);
Motion of a point on the semitrailer body
> `r_Pss^ss` := <x,y,z>;
> `r_Pss^so` := `R_ss^so`.`r_Pss^ss`;
> `v_O^so` := `R_to^so`.`v_O^to`;
> `v_Pss^so` := `v_O^so` + reldiff(`r_Pss^so`,`omega_so^so`);
> `a_Pss^so` := reldiff(`v_Pss^so`,`omega_so^so`);
Motion of a point on the hitch body
> `r_Ph^h` := <x,y,z>;
> `r_Ph^to` := `R_h^to`.`r_Ph^h`;
> `v_Ph^to` := `v_O^to` + reldiff(`r_Ph^to`,`omega_to^to`);
Wheel corner velocities
> `v_Wt^to` := `v_O^to` + reldiff(<x,y,z>,`omega_to^to`);
> `v_Ws^so` := `v_O^so` + reldiff(<x,y,z>,`omega_so^so`);
```

```
============ KINETICS ============
> `intF_t^to` := Multiply(`a_Pts^to`,dm);
> Map(s->
  expand(algsubs(dm=m_t,
  expand(algsubs(z*dm=rcg_tz*m_t,
  expand(algsubs(y*dm=rcg_ty*m_t,
  expand(algsubs(x*dm=rcg_tx*m_t,
  expand(s))))))))),`intF_t^to`);
> `intF_s^so` := Multiply(`a_Pss^so`,dm);
> Map(s->
  expand(algsubs(dm=m_s,
  expand(algsubs(z*dm=rcg_sz*m_s,
  expand(algsubs(y*dm=rcg_sy*m_s,
  expand(algsubs(x*dm=rcg_sx*m_s,
  expand(s))))))))),`intF_s^so`);
> `intM_t^to` := Multiply(CrossProduct(`r_Pts^to`,`a_Pts^to`),dm);
> Map(s->
  expand(algsubs(dm=m_t,
  expand(algsubs(z*dm=rcg_tz*m_t,
  expand(algsubs(y*dm=rcg_ty*m_t,
  expand(algsubs(x*dm=rcg_tx*m_t,
  expand(algsubs(y*z*dm=I_tyz,
  expand(algsubs(x*z*dm=I_txz,
  expand(algsubs(x*y*dm=I_txy,
  expand(algsubs(z^2*dm=I_tzz,
  expand(algsubs(y^2*dm=I_tyy,
  expand(algsubs(x^2*dm=I_txx,
  expand(s))))))))))))))))))),`intM_t^to`);
> `intM_s^so` := Multiply(CrossProduct(`r_Pss^so`,`a_Pss^so`),dm);
> Map(s->
  expand(algsubs(dm=m_s,
  expand(algsubs(z*dm=rcg_sz*m_s,
  expand(algsubs(y*dm=rcg_sy*m_s,
  expand(algsubs(x*dm=rcg_sx*m_s,
  expand(algsubs(y*z*dm=I_syz,
  expand(algsubs(x*z*dm=I_sxz,
  expand(algsubs(x*y*dm=I_sxy,
  expand(algsubs(z^2*dm=I_szz,
  expand(algsubs(y^2*dm=I_syy,
  expand(algsubs(x^2*dm=I_sxx,
  expand(s))))))))))))))))))),`intM_s^so`);
> `intF^to` := Add(`intF_t^to`,`R_so^to`.`intF_s^so`);
Hitch equations
> `m_s^ss` := <`m_sxy^ss`*cos(psi(t)),`m_sxy^ss`*sin(psi(t)),0>;
> `m_t^ts` := <Cc*(`phi_t`(t)-`phi_h`),`m_ty^ts`,`m_tz^ts`>;
> `M_h^h` := <`M_hx^h`,`M_hy^h`,`M_hz^h`>;
> eqnsh := map(s->s=0,convert(`R_ts^to`.`m_t^ts` +
  `R_so^to`.`R_ss^so`.`m_s^ss` + `R_h^to`.`M_h^h`,list));
> (A,b) := GenerateMatrix(eqnsh,[`m_sxy^ss`,`m_ty^ts`,`m_tz^ts`]);
============ DYNAMIC EQUATIONS ============
> subslist := [U(t)=U, V(t)=V, W(t)=W, Z(t)=Z, psi(t)=psi,
```

```
  diff(Z(t),t)=W, r(t)=r, diff(psi(t),t)=dpsi, psi(t)=psi,
  diff(‘phi_t‘(t),t)=‘dphi_t‘, ‘phi_t‘(t)=‘phi_t‘,
  diff(‘chi_t‘(t),t)=‘dchi_t‘, ‘chi_t‘(t)=‘chi_t‘,
  diff(‘phi_s‘(t),t)=‘dphi_s‘, ‘phi_s‘(t)=‘phi_s‘,
  diff(‘chi_s‘(t),t)=‘dchi_s‘, ‘chi_s‘(t)=‘chi_s‘,
  diff(U(t),t)=dU, diff(V(t),t)=dV, diff(W(t),t)=dW,
  diff(Z(t),t,t)=dW, diff(r(t),t)=dr, diff(psi(t),t,t)=ddpsi,
  diff(‘phi_t‘(t),t,t) = ‘ddphi_t‘, diff(‘chi_t‘(t),t,t)=‘ddchi_t‘,
  diff(‘phi_s‘(t),t,t)=‘ddphi_s‘, diff(‘chi_s‘(t),t,t)=‘ddchi_s‘];
> intM_t := map(s->subs(subslist,s), ‘intM_t^to‘);
> intM_s := map(s->subs(subslist,s), ‘intM_s^so‘);
> intF := map(s->subs(subslist,s), ‘intF^to‘);
> dstates := [dU, dV, dW, dr, ddpsi, ‘ddphi_t‘, ‘ddchi_t‘,
  ‘ddphi_s‘, ‘ddchi_s‘];
> (H,F) := GenerateMatrix(map(s->s=0,[op(convert(intF,list)),
  op(convert(intM_t,list)), op(convert(intM_s,list))]),dstates);
============ C-CODE GENERATION ============
> with(codegen,C);
> C(map(s->subs(subslist,s),‘r_Pts^to‘),filename="r_Ptsto_maple.c");
> C(map(s->subs(subslist,s),‘v_Pts^to‘),filename="v_Ptsto_maple.c");
> C(map(s->subs(subslist,s),‘a_Pts^to‘),filename="a_Ptsto_maple.c");
> C(map(s->subs(subslist,s),‘r_Pss^so‘),filename="r_Pssso_maple.c");
> C(map(s->subs(subslist,s),‘v_Pss^so‘),filename="v_Pssso_maple.c");
> C(map(s->subs(subslist,s),‘a_Pss^so‘),filename="a_Pssso_maple.c");
> C(map(s->subs(subslist,s),‘r_Ph^to‘),filename="r_Phto_maple.c");
> C(map(s->subs(subslist,s),‘v_Ph^to‘),filename="v_Phto_maple.c");
> C(map(s->subs(subslist,s),‘R_so^to‘.‘r_Pss^so‘),filename="r_Pssto_maple.c");
> C(H,filename="H_maple.c");
> C(F,filename="F_maple.c");
> C(map(s->subs(subslist,s),‘v_Wt^to‘),filename="v_Ptto_maple.c");
> C(map(s->subs(subslist,s),‘v_Ws^so‘),filename="v_Psso_maple.c");
```

D. Model Parameters

Chassis	
m_{ts}	Tractor sprung mass
m_{ss}	Semitrailer sprung mass
$\bar{r}_{CG,ts}$	Tractor sprung body CM location
$\bar{r}_{CG,ss}$	Semitrailer sprung body CM location
$\underline{I}_{ts}$	Tractor sprung body inertia matrix
$\underline{I}_{ss}$	Semitrailer sprung body inertia matrix
C_c	Tractor frame torsional stiffness

Axles and suspensions	
$\bar{r}_{w,l}$	Left wheel location
$\bar{r}_{w,r}$	Right wheel location
$\bar{r}_{s,l}$	Left suspension location
$\bar{r}_{s,r}$	Right suspension location
C_s	Suspension vertical stiffness
D_s	Suspension vertical damping coefficient
C_r	Roll-bar stiffness
h_s	Unloaded suspension length
C_{steer}	Steering compliance
K_{steer}	Geometric steering-ratio
K_{roll}	Roll-steer coefficient

Wheels and tires	
J_w	Wheel inertia
R_w	Wheel radius
$\mu_x(\lambda)$	Longitudinal adhesion coefficient mapping at pure braking
$\mu_y(\alpha)$	Lateral adhesion coefficient mapping at pure cornering

Paper II

A Novel Semi-Empirical Tire Model for Combined Slips

Magnus Gäfvert Jacob Svendenius

Abstract

A new tire-force model for simultaneous braking and cornering is presented, which is based on combining empirical models for pure braking and cornering with brush-model tire mechanics. The aim is to offer an easy-to-use, accurate model for vehicle-handling simulations. On a working tire the contact patch between the tire and the road is, in general, divided into an adhesion region where the rubber is gripping the road and a sliding region where the rubber slides on the road surface. The total force generated by the tire is then composed of components from these two regions. The brush model describes this in a mechanical framework. The proposed model is based on a new method to extract adhesion and sliding forces from empirical pure-slip tire models. These forces are then scaled to account for the combined-slip condition. The combined-slip self-aligning torque is also described. A particular feature of the model is the inclusion of velocity dependence, even if this is not explicitly present in the empirical pure-slip model. The approach is quite different from most previous combined-slip models, in that it is based on a rather detailed mechanical model in combination with empirical pure-slip models. The model is computationally sound and efficient and does not rely on any additional parameters that depend on combined-slip data for calibration. It can be used in combination with virtually any empirical pure-slip model and in this work the Magic Formula is used in examples. Results show good correspondence with experimental data.

1. Introduction

The forces generated in the contact between the tires and the road are of major importance for the dynamic behavior of a road vehicle. Hence, accurate tire models are necessary components of complete vehicle models aimed at analyzing or simulating vehicle motion in real driving conditions. There are many previous models which describe the tire-forces generated at the pure-slip conditions of braking, driving, or cornering, as well as models which describe the tire forces resulting from combined-slip conditions of simultaneous braking (or driving) and cornering. Some models are theoretical in the sense that they aim at modeling the physical processes that generate the forces. Other models are empirically oriented and aim at describing observed phenomena in a simple form. Theoretical models are, generally, based on simplifying assumptions, which limit their practical use. They often involve parameters which are difficult to identify. Empirical models are, in general, based on functional approximations of experimental data. Tire forces for a specific surface and tire type depend on slip ratio, slip angle, normal load, and wheel-travel velocity. It is impractical to collect and manage measurement data for all possible combinations in these four dimensions. Models which reproduce tire behavior from data of smaller dimensionality are therefore of great help. Pure-slip tire forces at varying normal load may be described well by compact empirical models, which are widely accepted. Tire data for calibration of those models are straightforward to measure in test-bench experiments. At combined-slip conditions, functional approximation with ad-hoc formulas is more difficult, due to the higher dimensionality. Empirical models tend to be either rough approximations or quite complex and difficult to understand. Accurate models rely on parameters that need to be calibrated with combined-slip experimental data.

The main result in this work is the derivation of a new model for combined-slip conditions. The proposed model is based on available empirical pure-slip models and does not rely on combined-slip data for calibration. When a tire generates contact force it is, in general, partly gripping and partly sliding on the road surface. The idea behind the model is to separate the empirical pure-slip forces into components of adhesion and sliding, which are then used to compute the combined-slip forces. The procedures for this are based on the standard brush-model. The proposed model has several attractive features:

- It reduces exactly to the empirical model at pure-slip.
- It gives a smooth transition from small-slip to large-slip behavior that agrees with empirical observations.

- Only few parameters are needed, which all have clear physical interpretations.
- Nominal parameter values may be derived automatically from the empirical pure-slip models.
- Differences between driving and braking conditions are accounted for.
- Velocity dependence is included, even when using velocity-independent empirical pure-slip models.
- The model integrates smoothly with Magic Formula pure-slip models.

In the remaining part of this section necessary definitions are introduced. In Section 2 the brush model is briefly recapitulated. Section 3 presents the procedure to decompose empirical pure-slip tire forces into components of adhesion and sliding. It is described how combined-slip forces may be constructed from these components. In Section 4 this is used to formulate the semi-empirical model. Practical implementation of the model is discussed. Results and validation of the model are then presented and discussed in Section 5. The paper concludes with a summary in Section 6.

Tire Kinematics

This section describes the relevant tire kinematics and introduces definitions that are used in the following. Refer to Figure 1 for illustration. Vectors have two components and are denoted by a bar as in $\bar{u}$. The corresponding components and magnitude are denoted by u_x, u_y, and u.

The *wheel-travel velocity* $\bar{v} = (v_x, v_y)$ deviates from the wheel heading by the slip angle α:

$$\tan(\alpha) = \frac{v_y}{v_x} \tag{1}$$

The circumferential velocity of the wheel is

$$v_c = \Omega R_e \tag{2}$$

where Ω is the wheel angular velocity, and R_e the effective rolling-radius of the tire. The *slip velocity*, or the relative motion of the tire in the contact patch to ground, is

$$\bar{v}_s = (v_x - v_c, v_y) \tag{3}$$

The direction of the slip velocity is denoted by

$$\tan(\beta) = \frac{v_{sy}}{v_{sx}} \tag{4}$$

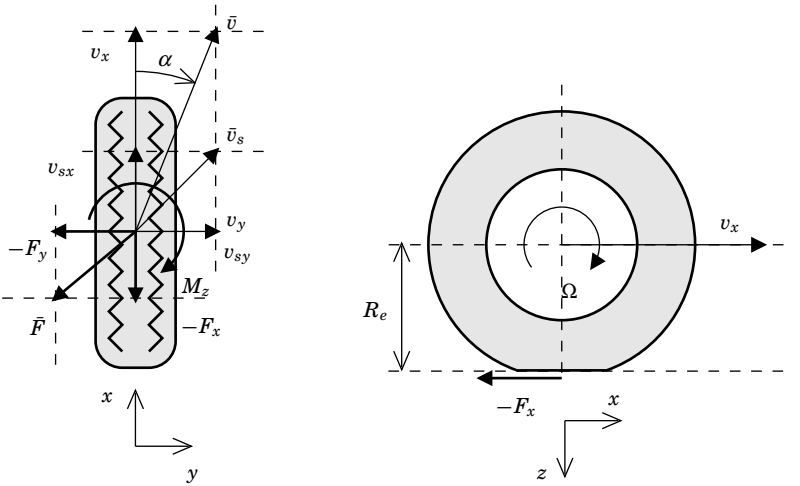

Figure 1 Kinematics of a tire during braking and cornering. Force vectors are also shown. (Left: top view; Right: side view)

The *tire slip* is defined by normalizing the slip-velocity with a reference velocity. Three slip definitions are commonly used:

$$\bar{\sigma} = \frac{\bar{v}_s}{v_c} \tag{5a}$$

$$\bar{\kappa} = \frac{\bar{v}_s}{v_x} \tag{5b}$$

$$\bar{s} = \frac{\bar{v}_s}{v} \tag{5c}$$

Note that the slips are collinear with the slip velocity $\bar{v}_s$. It is the custom to describe tire-forces as functions of the slip, rather than the slip velocity. This convention is followed also in this work. Implicitly, this assumes that the forces do not depend on the magnitude of the slip velocity, v_s. In general, at least the sliding friction is velocity dependent. There are several conventions on how to define the tire slips, e.g. the ISO and SAE standards [ISO 8855, 1991, SAE Recommended Practice J670e, 1976] use $-100\kappa_x$ [%] to represent longitudinal slip, and α [deg] for lateral slip. In this paper, the slips are defined such that signs are consistent for the different slip definitions and such that a generated tire force has opposite sign to the slip. This means that braking or left cornering will result in positive slip and negative force. For convenience, the *slip ratio*, λ, is used

to denote longitudinal slip as

$$\lambda = \kappa_x \tag{6}$$

It is straightforward to translate between the different slip representations:

$$\bar{\sigma} = (\lambda, \tan(\alpha)) / (1 - \lambda) = \frac{\bar{\kappa}}{1 - \kappa_x} = \frac{\bar{s}}{\sqrt{1 - s_y^2} - s_x} \tag{7a}$$

$$\bar{\kappa} = (\lambda, \tan(\alpha)) = \frac{\bar{\sigma}}{1 + \sigma_x} = \frac{\bar{s}}{\sqrt{1 - s_y^2}} \tag{7b}$$

$$\bar{s} = (\lambda \cos(\alpha), \sin(\alpha)) = \frac{\bar{\sigma}}{\sqrt{(1 + \sigma_x)^2 + \sigma_y^2}} = \frac{\bar{\kappa}}{\sqrt{1 + \kappa_y^2}} \tag{7c}$$

2. The Brush Model

The brush model is a standard approach to model tire forces, see e.g. [Pacejka, 1988, Gim and Nikravesh, 1991, Wong, 2001, Gäfvert and Svendenius, 2003]. Therefore, this section is kept brief and has the main purpose of introducing notation and formulas which are needed in the succeeding sections.

The brush model describes the generation of tire forces by dividing the contact patch into regions of adhesion and sliding. The rubber volume between the tire and the road surface is partitioned into infinitesimal elements in the form of elastic rectangular blades, or bristles, which stretches laterally over the entire contact region, see Figure 2. Each bristle is assumed to deform independently and linearly elastic in the longitudinal and lateral directions. In the adhesion region the bristles adhere to the road surface. Thus, the deformation force is carried by static friction. In the sliding region the bristles slide on the road surface under influence of kinetic friction. Positions in the contact region are expressed in a reference system attached to the carcass, with the origin located in the center of the contact region. The length of the contact region is $2a$.

The carcass is assumed to be stiff, which means that effects of carcass deformation are neglected. The bristle deformation is regarded in quasistatic condition.

Combined-Slip Adhesion Forces

Bristle Forces. Regard the specific infinitesimal bristle which is attached to the carcass at position x, relative an origo in the center of the

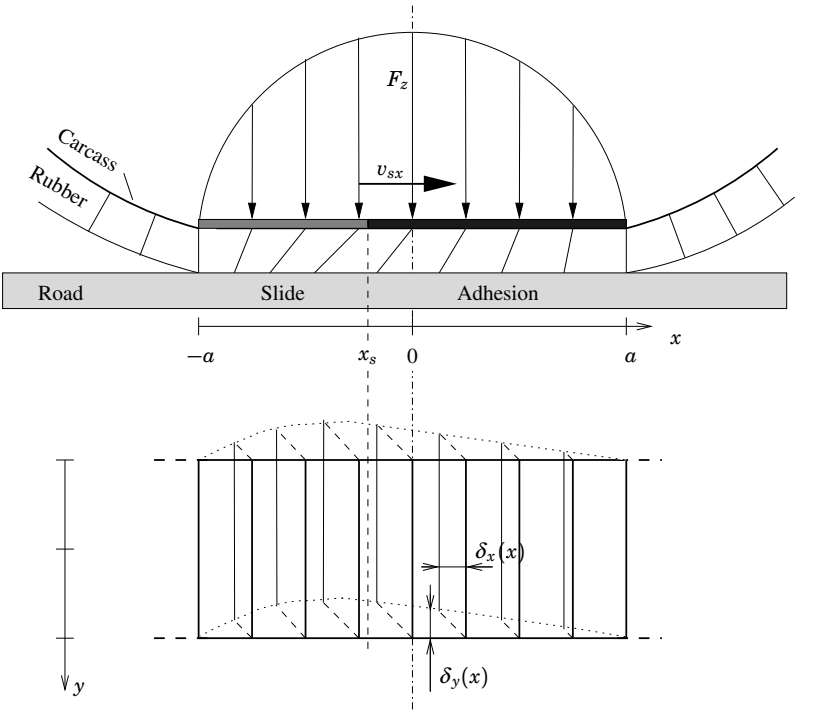

Figure 2 The deformation of the rubber layer between the tire carcass and the road, according to the brush model. The carcass moves with the slip velocity, $\bar{v}_s$, relative the road. The contact zone moves with the wheel-travel velocity, $\bar{v}$. (Top: side view; Bottom: top view)

contact patch. Assume that this bristle belongs to the adhesion region. The bristle is in contact with the road surface at position $x_r(x)$, $y_r(x)$. The contact-point position may be described by

$$x_r(x) = a - \int_0^{t_c(x)} v_x \, dt \tag{8a}$$

$$y_r(x) = -\int_0^{t_c(x)} v_y \, dt \tag{8b}$$

where $t_c(x)$ is the time elapsed since the bristle entered the contact region. The velocities v_c, v_x, and v_y are assumed to be constant during the integration interval $[0, t_c(x)]$, i.e., as a bristle travels through the adhesion region. Hence, the bristle position is $x = a - v_c t_c(x)$ and $t_c(x) = (a-x)/v_c$.

Using (8) and the expressions for $t_c(x)$, the deformation of the bristle is

$$\delta_x(x) = x_r(x) - x = -\frac{v_x - v_c}{v_c}(a - x) = -\sigma_x (a - x) \tag{9a}$$

$$\delta_y(x) = y_r(x) = -\frac{v_y}{v_c}(a - x) = -\sigma_y (a - x) \tag{9b}$$

where the slip definition (5) is used in the last equality. With the assumption on linear elasticity, the deformation forces corresponding to (9) is

$$dF_{ax}(x) = c_{px}\, dx\, \delta_x(x) \tag{10a}$$

$$dF_{ay}(x) = c_{py}\, dx\, \delta_y(x) \tag{10b}$$

where c_{px} and c_{py} are the longitudinal and lateral bristle stiffnesses per unit length. The assumption of constant v_c, v_x, v_y in the interval $[0, t_c(x)]$ is relaxed to slow variations in σ_x and σ_y, with respect to the duration $2a/v_c$, which is the time for a bristle to travel through the contact patch. The total adhesion force is computed by integration of (10) over the adhesion region. With (9), this gives

$$F_{ax} = \int_{x_s}^{a} dF_{ax}(x) = -c_{px}\sigma_x \int_{x_s}^{a} (a - x)\, dx \tag{11a}$$

$$F_{ay} = \int_{x_s}^{a} dF_{ay}(x) = -c_{py}\sigma_y \int_{x_s}^{a} (a - x)\, dx \tag{11b}$$

The break-away point, x_s, is the point which separates the regions of adhesion and sliding. To compute the total adhesion force it is necessary to know the location of x_s.

The Size of the Adhesion Region. The size of the adhesion region is determined by the available static friction between the tire and the road. The deformation is limited by the largest force that can be carried by the friction. The static friction is assumed to be anisotropic, with the friction coefficients μ_{sx} and μ_{sy}, respectively. With a normal force $dF_z(x)$ acting on the infinitesimal bristle at position x, the available static friction force is described by the elliptic constraint

$$\left(\frac{dF_{ax}(x)}{dF_z(x)\mu_{sx}}\right)^2 + \left(\frac{dF_{ay}(x)}{dF_z(x)\mu_{sy}}\right)^2 \leq 1 \tag{12}$$

As a result, the magnitude of the available friction force is dependent on the direction of the deformation force $d\bar{F}_a(x)$, given by (10). When $d\bar{F}_a(x)$

exceeds the friction constraint, the bristle leaves the adhesion region and starts to slide. Assume a parabolic pressure distribution

$$q_z(x) = \frac{3F_z}{4a}\left(1 - \left(\frac{x}{a}\right)^2\right) \tag{13}$$

with $dF_z(x) = q_z(x)\,dx$. At this point, the following definitions are introduced to simplify notation.

Definition 1 (Limit slips)**.** Define the limit slips as

$$\sigma_x^\circ \triangleq \frac{3F_z\mu_{sx}}{2a^2c_{px}} \tag{14a}$$

$$\sigma_y^\circ \triangleq \frac{3F_z\mu_{sy}}{2a^2c_{py}} \tag{14b}$$

Definition 2 (Normalized slip)**.** Define the normalized slip as

$$\psi(\sigma_x,\sigma_y) \triangleq \sqrt{\left(\frac{\sigma_x}{\sigma_x^\circ}\right)^2 + \left(\frac{\sigma_x}{\sigma_x^\circ}\right)^2} \tag{15}$$

and the normalized-slip angle as

$$\tan(\beta^\circ) \triangleq \frac{\psi(0,\sigma_y)}{\psi(\sigma_x,0)} \tag{16}$$

Now, combining (9), (10), and (13) with (12) and solving for the breakaway point x_s, gives

$$x_s(\sigma_x,\sigma_y) = (2\psi(\sigma_x,\sigma_y) - 1)a \tag{17}$$

It is clear that partial sliding occurs when $-a < x < a$, i.e., when $\psi(\sigma_x,\sigma_y) < 1$. At full sliding, $\psi(\sigma_x,\sigma_y) \geq 1$, there is no adhesion region and $F_{ax}(\sigma_x,\sigma_y) = F_{ay}(\sigma_x,\sigma_y) = 0$. In the case of pure slip, when either σ_x or σ_y is zero, then full sliding occurs at the limit slips $\sigma_x = \sigma_x^\circ$ or $\sigma_y = \sigma_y^\circ$.

Total Adhesion Force. The adhesion forces are obtained by solving the integrals (11), using the size of the adhesion region as given by (17).

Property 1 (Adhesion forces)**.** *At partial sliding ($\psi(\sigma_x,\sigma_y) < 1$), the adhesion forces are*

$$F_{ax}(\sigma_x,\sigma_y) = -2a^2c_{px}\sigma_x\left(1 - \psi(\sigma_x,\sigma_y)\right)^2 \tag{18a}$$

$$F_{ay}(\sigma_x,\sigma_y) = -2a^2c_{py}\sigma_y\left(1 - \psi(\sigma_x,\sigma_y)\right)^2 \tag{18b}$$

and at full sliding ($\psi(\sigma_x,\sigma_y) \geq 1$) then $F_{ax}(\sigma_x,\sigma_y) = F_{ay}(\sigma_x,\sigma_y) = 0$.

Special notations for the forces at pure slip are introduced as

$$F_{0ax}(\sigma_x) \triangleq F_{ax}(\sigma_x, 0) \tag{19a}$$

$$F_{0ay}(\sigma_y) \triangleq F_{ay}(0, \sigma_y) \tag{19b}$$

Note that it follows from (10) and (9) that the longitudinal and lateral adhesion forces per unit length in the contact region are produced independently. The effect of combined slips on the total adhesion force is solely due to the variations in size of the adhesion region.

Combined-Slip Sliding Forces

The normal force acting on the sliding region at partial sliding may be computed from (13) and (17) as

$$F_{sz}(\sigma_x, \sigma_y) = \int_{-a}^{x_s(\sigma_x,\sigma_y)} q_z(x)\, dx = F_z \psi^2(\sigma_x, \sigma_y)\,(3 - 2\psi(\sigma_x, \sigma_y)) \tag{20}$$

At full sliding then $F_{sz}(\sigma_x, \sigma_y) = F_z$. The sliding forces are assumed to be produced by isotropic kinetic friction, as $\bar{F}_{sx}(\sigma_x, \sigma_y) = -\mu_{\mathrm{k}} F_{sz}(\sigma_x, \sigma_y)\bar{v}_s / v_s$. Hence, the resulting force is collinear with the slip velocity $\bar{v}_s$ and, consequently, with the slip.

Even if the friction is supposed to be isotropic, it may be the case that different friction coefficients $\mu_{\mathrm{k}x}$ and $\mu_{\mathrm{k}y}$ are given longitudinally and laterally. A reason for this could, for example, be measurement errors. A collinear friction force is still obtained with:

Property 2 (Sliding forces)**.** *Collinear sliding forces are described by*

$$F_{sx}(\sigma_x, \sigma_y) = -\cos(\beta')\mu_{\mathrm{k}x} F_{sz}(\sigma_x, \sigma_y) \tag{21a}$$

$$F_{sy}(\sigma_x, \sigma_y) = -\sin(\beta')\mu_{\mathrm{k}y} F_{sz}(\sigma_x, \sigma_y) \tag{21b}$$

where β' is defined as

$$\tan(\beta') \triangleq \left(\frac{\mu_{\mathrm{k}y}}{\mu_{\mathrm{k}x}}\right)^{-1} \frac{v_{sy}}{v_{sx}} = \left(\frac{\mu_{\mathrm{k}y}}{\mu_{\mathrm{k}x}}\right)^{-1} \tan(\beta) \tag{22}$$

The resulting force acts in the opposite direction to the sliding motion. The friction coefficient depends on the slip-velocity angle, β, and vary between $\mu_{\mathrm{k}x}$ and $\mu_{\mathrm{k}y}$. In the special case of pure slip, the sliding forces are

$$F_{0sx}(\sigma_x) = -\mu_{\mathrm{k}x} F_{sz}(\sigma_x, 0)\,\mathrm{sgn}(\sigma_x) \tag{23a}$$

$$F_{0sy}(\sigma_y) = -\mu_{\mathrm{k}y} F_{sz}(0, \sigma_y)\,\mathrm{sgn}(\sigma_y) \tag{23b}$$

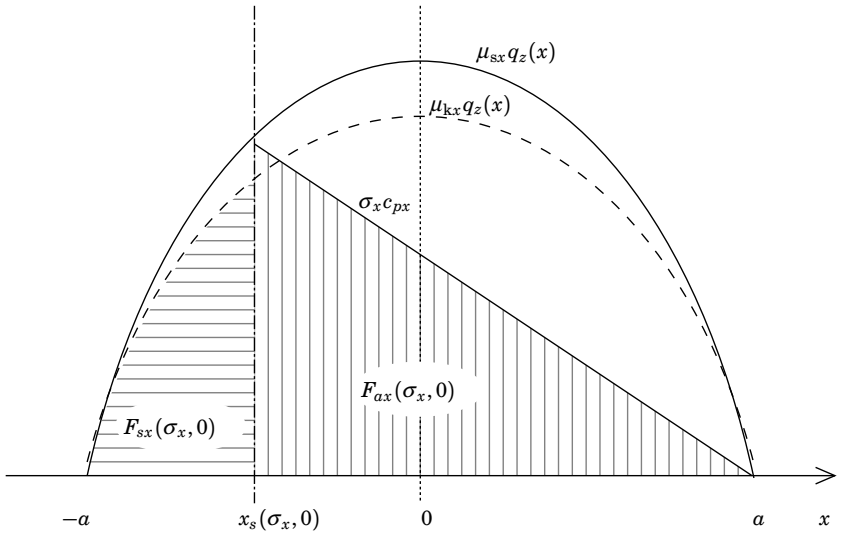

Figure 3 Illustration of partition of the contact area into regions of adhesion and sliding, for the case of pure longitudinal slip. The adhesion force for an element at x is $c_{px}\sigma_x(a-x)\,dx$. The sliding force for an element at x is determined by the pressure distribution, as $\mu_{kx} q_z(x)\,dx$. The vertically and horizontally striped areas represent the total forces of adhesion and sliding.

Combined-Slip Total Forces

The total tire force is given by adding the adhesion forces of (18) and the sliding forces of (21):

$$F_x(\sigma_x, \sigma_y) = F_{ax}(\sigma_x, \sigma_y) + F_{sx}(\sigma_x, \sigma_y) \tag{24a}$$

$$F_y(\sigma_x, \sigma_y) = F_{ay}(\sigma_x, \sigma_y) + F_{sy}(\sigma_x, \sigma_y) \tag{24b}$$

In Figure 3, the generation of forces in the regions of adhesion and sliding is illustrated for the case of a pure longitudinal slip, $(\sigma_x, 0)$. The vertically striped area corresponds to the adhesion force and the horizontally striped area to the sliding force. The adhesion force grows linearly with the slope $c_{px}\sigma_x$, as a contact element moves into the adhesion region. The slope corresponding to full sliding, i.e., when $\sigma_x = \sigma_x^\circ$, is also shown. The sliding force per unit length is $\mu_{kx} q_z(x)$, since $q_z(x)$ is the normal force per unit length. The case of a pure lateral slip is analogous. The effect of a combined slip, (σ_x, σ_y), on the longitudinal force is illustrated in Figure 4. From (17) it is clear that the adhesion region shrinks, compared to the case with pure slip. The sliding region, thus, grows accordingly. The produced adhesion force per unit length is the same as for the pure-slip case.

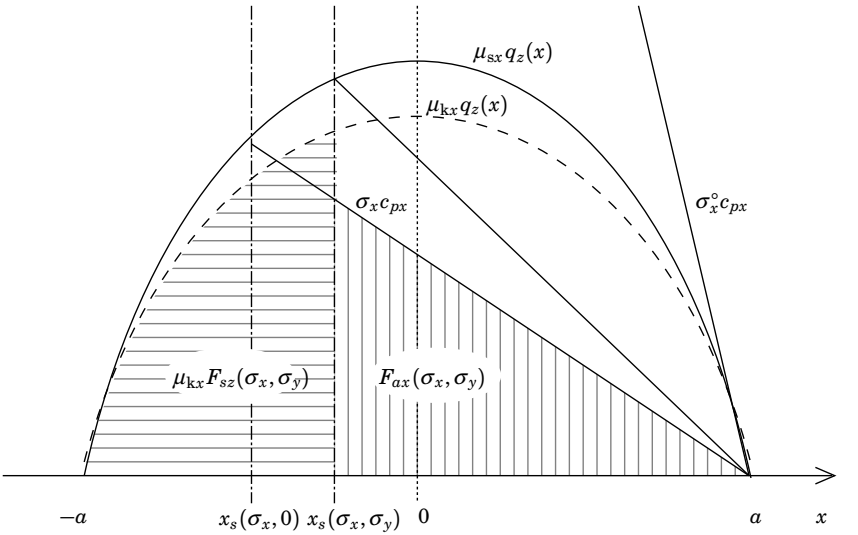

Figure 4 Illustration of the effect of combined slip. The combined-slip has the effect of decreasing the size of the adhesion region, compare with Figure 3.

Hence, the slope is the same but the area corresponding to the force is smaller, since the adhesion region is smaller. Note that the horizontally striped area corresponds to $\mu_{kx} F_{sz}(\sigma_x, \sigma_y)$, which is the force that would result for pure longitudinal sliding with the break-away point $x_s(\sigma_x, \sigma_y)$. This force is longitudinally constrained by friction, according to (21). The effect on the lateral force is analogous.

The braking and cornering stiffnesses are the linearizations of the pure-slip friction curves at small slips and may be computed from (24):

$$C_x \triangleq -\left.\frac{\partial F_x(\sigma_x, 0)}{\partial \sigma_x}\right|_{\sigma_x=0} = 2c_{px}a^2 \tag{25a}$$

$$C_y \triangleq -\left.\frac{\partial F_y(0, \sigma_y)}{\partial \sigma_y}\right|_{\sigma_y=0} = 2c_{py}a^2 \tag{25b}$$

Self-Aligning Torque

The self-aligning torque consists of two parts. The main part, M_z', is the torque developed by the non-symmetric distribution of the lateral force F_y. An additional part, M_z'', comes up because of the deformation of the

tire. The torque dM_z' developed at position x in the contact region is

$$dM_z'(x) = dF_y(x)\, x \tag{26}$$

In the adhesion part of the contact region, the expression for $dF_y(x)$ is given by (10), together with (9). In the sliding zone, it is instead given by differentiating (21), using $dF_z(x) = q(x)\, dx$. Integrating over the regions of adhesion and sliding separately gives

$$\begin{aligned} M_{az}'(\sigma_x, \sigma_y) &= \int_{x_s(\sigma_x,\sigma_y)}^{a} dF_{ay}(x)\, x \\ &= -c_{py} a^3 \sigma_y \frac{2}{3} (1 - \psi(\sigma_x, \sigma_y))^2 (4\psi(\sigma_x, \sigma_y) - 1) \end{aligned} \tag{27}$$

$$\begin{aligned} M_{sz}'(\sigma_x, \sigma_y) &= \int_{-a}^{x_s(\sigma_x,\sigma_y)} dF_{sy}(x)\, x \\ &= -3\mu_{\mathrm{k}x} \sin(\beta) a F_z \psi^2(\sigma_x, \sigma_y)(1 - \psi(\sigma_x, \sigma_y))^2 \end{aligned} \tag{28}$$

and

$$M_z'(\sigma_x, \sigma_y) = M_{az}'(\sigma_x, \sigma_y) + M_{sz}'(\sigma_x, \sigma_y) \tag{29}$$

When there is a lateral slip, the tire deflects laterally and the point of action for the longitudinal force will have an offset from the central plane of the wheel. This produces an additional deformation-torque in the z-direction. A longitudinal deflection, together with a lateral force, has the same effect. Since it is assumed that the carcass is stiff, the deformation is here described by bristle deflections. The deformation torque developed at position x in the contact region is then described by

$$dM_z''(x) = dF_y(x)\delta_x(x) - dF_x(x)\delta_y(x) \tag{30}$$

Like above, integration over the regions is performed separately. The deformation $\delta_x(x)$ is computed from (9) in the adhesion region and from (10), using the infinitesimal sliding force, in the sliding region. Hence,

$$\begin{aligned} M_{az}''(\sigma_x, \sigma_y) &= \int_{x_s(\sigma_x,\sigma_y)}^{a} dF_{ay}(x)\delta_x(x) - dF_{ax}(x)\delta_y(x) \\ &= \frac{4}{3}\left(\frac{1}{C_x} - \frac{1}{C_y}\right) \frac{a}{(1 - \psi(\sigma_x, \sigma_y))} F_{ax}(\sigma_x, \sigma_y) F_{ay}(\sigma_x, \sigma_y) \end{aligned} \tag{31}$$

$$M''_{sz}(\sigma_x,\sigma_y) = \int_{-a}^{x_s(\sigma_x,\sigma_y)} dF_{sy}(x)\delta_x(x) - dF_{sx}(x)\delta_y(x)$$
$$= \frac{6}{5}\left(\frac{1}{C_x} - \frac{1}{C_y}\right)\frac{a(10 - 15\psi(\sigma_x,\sigma_y) + 6\psi^2(\sigma_x,\sigma_y))}{\psi(\sigma_x,\sigma_y)(3 - 2\psi(\sigma_x,\sigma_y))^2}$$
$$\cdot F_{sx}(\sigma_x,\sigma_y)F_{sy}(\sigma_x,\sigma_y) \quad (32)$$

where (18) and (21) are used in the last step. Finally,

$$M''_z(\sigma_x,\sigma_y) = M''_{az}(\sigma_x,\sigma_y) + M''_{sz}(\sigma_x,\sigma_y) \quad (33)$$

and

$$M_z(\sigma_x,\sigma_y) = M'_z(\sigma_x,\sigma_y) + M''_z(\sigma_x,\sigma_y) \quad (34)$$

The pneumatic trail denotes the distance between the center of the tire and point of action for the lateral force. It is defined as $t(\sigma_x,\sigma_y) = M_z(\sigma_x,\sigma_y)/F_y(\sigma_x,\sigma_y)$. The adhesion trail is then $t_a(\sigma_x,\sigma_y) = M_{az}(\sigma_x,\sigma_y)/F_{ay}(\sigma_x,\sigma_y)$ and the sliding trail $t_s(\sigma_x,\sigma_y) = M_{sz}(\sigma_x,\sigma_y)/F_{sy}(\sigma_x,\sigma_y)$. By using (27) and (18), respective (28) and (21), the contributions from $M'_z(\sigma_x,\sigma_y)$ to the pneumatic trail, $t'_a(\sigma_x,\sigma_y)$ and $t'_s(\sigma_x,\sigma_y)$, are given by

$$t'_a(\sigma_x,\sigma_y) = \frac{a}{3}(4\psi(\sigma_x,\sigma_y) - 1) \qquad t'_s(\sigma_x,\sigma_y) = -3a\frac{(1-\psi(\sigma_x,\sigma_y))^2}{(3-2\psi(\sigma_x,\sigma_y))} \quad (35)$$

The contributions to the trails from from $M''_z(\sigma_x,\sigma_y)$ can be identified directly from (31) and (32).

In analogy with the braking and cornering stiffnesses, the aligning stiffness is defined as

$$C_z \triangleq -\left.\frac{\partial M_z}{\partial \sigma_y}\right|_{\sigma_x,\sigma_y=0} = c_{py}a^3\frac{2}{3} = C_y\frac{a}{3} \quad (36)$$

3. Brush-Model Relations Between Pure- and Combined-Slip Conditions

In this section a number of properties are stated, which describe relations between pure-slip and combined-slip forces for the brush model. In Section 4 they are combined with empirical data, to form a semi-empirical tire-force model for combined slips. The basic idea is to treat sliding and adhesion forces separately. Since empirical pure-slip data normally only regard total tire-forces, it is necessary to have a procedure to separate the forces from the regions of adhesion and sliding and adhesive.

Separation of Adhesion Forces and Sliding Forces

Property 3 (Separation of Adhesion Forces and Sliding Forces)**.** *The adhesion part of the total tire-force is given by*

$$F_{0ax}(\sigma_x) = \big(1 - \theta(\psi(\sigma_x, 0))\big) F_{0x}(\sigma_x) \tag{37a}$$

$$F_{0ay}(\sigma_y) = \big(1 - \theta(\psi(0, \sigma_y))\big) F_{0y}(\sigma_y) \tag{37b}$$

and the sliding part by

$$F_{0sx}(\sigma_x) = \theta(\psi(\sigma_x, 0)) F_{0x}(\sigma_x) \tag{38a}$$

$$F_{0sy}(\sigma_y) = \theta(\psi(0, \sigma_y)) F_{0y}(\sigma_y) \tag{38b}$$

with

$$\theta(\varphi) = \frac{\varphi\,(3 - 2\varphi)}{3\rho\,(1 - \varphi)^2 + \varphi\,(3 - 2\varphi)} \tag{39}$$

for $\varphi < 1$ *and* $\theta(\varphi) = 1$ *for* $\varphi \geq 1$*. The parameter* $\rho \triangleq \mu_{sx}/\mu_{kx} = \mu_{sy}/\mu_{ky} > 0$ *describes the ratio between the static and kinetic friction coefficients and is assumed to be equal in the longitudinal and lateral directions.*

Proof. See Appendix A. □

Remark 1. Note that $\theta(\varphi)$ and $1 - \theta(\varphi)$ of (39) are non-singular for all $0 \leq \varphi \leq 1$, since the denominator polynomial $3\rho(1-\varphi)^2 + \varphi(3 - 2\varphi)$ does not have any real roots in the interval $[0, 1]$ for $\rho > 0$.

Pure-Slip Mappings

The aim is to describe combined-slip forces in terms of corresponding pure-slip forces. Then, it needs to be determined which pure slips that correspond to a particular combined slip. The bristle deformations are the source of the adhesion forces. Therefore, for adhesion forces it makes sense to regard pure slips which result in the same deformation as the combined slip.

Property 4 (Deformation-invariant pure slips)**.** *The pure slips* $(\sigma_x, 0)$ *and* $(0, \sigma_y)$ *result in the same longitudinal and lateral bristle deformation, respectively, as the combined slip* (σ_x, σ_y) *at partial sliding* $(\psi(\sigma_x, \sigma_y) < 1)$*.*

Proof. See Appendix A. □

Note that the deformation-invariant pure slips only have the physical interpretation of deformation in the adhesion region. Kinetic friction is the source of the sliding forces. Since kinetic friction, in general, is velocity

dependent it makes sense to regard pure slips that result in the same slip velocity as the combined slip.

Property 5 (Slip-velocity invariant pure-slips)**.** *The pure slips* $(\sigma_{0x}^{\text{vel}}, 0)$ *and* $(0, \sigma_{0y}^{\text{vel}})$ *at the wheel-travel velocity* v_0*, with*

$$\sigma_{0x}^{\text{vel}} = \frac{\rho_v \sqrt{\sigma_x^2 + \sigma_y^2}\, \text{sgn}(\sigma_x)}{\sqrt{(1+\sigma_x)^2 + \sigma_y^2} - \rho_v \sqrt{\sigma_x^2 + \sigma_y^2}\, \text{sgn}\,(\sigma_x)} \tag{40a}$$

$$\sigma_{0y}^{\text{vel}} = \frac{\rho_v \sqrt{\sigma_x^2 + \sigma_y^2}\, \text{sgn}(\sigma_y)}{\sqrt{\left((1+\sigma_x)^2 + \sigma_y^2\right) - \rho_v^2(\sigma_x^2 + \sigma_y^2)}} \tag{40b}$$

where $\rho_v = v/v_0$*, result in the same slip velocity,* v_s*, as the combined slip* (σ_x, σ_y) *at the wheel-travel velocity* v*.*

Proof. See Appendix A. □

It can also be argued that both the adhesion force and the sliding force depend on the sizes of the adhesion and sliding regions. Then, it makes sense to regard the pure slips that result in the same size as the combined slip.

Property 6 (Region-invariant pure slips)**.** *The pure slips* $(\sigma_{0x}^{\text{reg}}, 0)$ *and* $(0, \sigma_{0y}^{\text{reg}})$*, with*

$$\sigma_{0x}^{\text{reg}} = \sigma_x^{\circ} \psi(\sigma_x, \sigma_y)\, \text{sgn}(\sigma_x) \tag{41a}$$

$$\sigma_{0y}^{\text{reg}} = \sigma_y^{\circ} \psi(\sigma_x, \sigma_y)\, \text{sgn}(\sigma_y) \tag{41b}$$

result in adhesion and sliding regions of the same size as the combined slip (σ_x, σ_y)*.*

Proof. See Appendix A. □

The region-invariant pure slips do not have clear physical interpretations in case of full sliding.

Combined-Slip Forces as Scaling of Pure-Slip Forces

In the following properties it is stated that combined-slip forces can be described in terms of scaling the corresponding pure-slip forces, for the pure slips defined in the previous section.

Property 7 (Deformation-invariant adhesion forces)**.** *Combined-slip adhesion forces at partial sliding ($\psi(\sigma_x, \sigma_y) < 1$) are related to deformation-invariant pure-slip forces as:*

$$F_{ax}(\sigma_x, \sigma_y) = \frac{3\rho\left(1 - \psi(\sigma_x, \sigma_y)\right)^2}{3\rho\left(1 - \psi(\sigma_x, 0)\right)^2 + \psi(\sigma_x, 0)\left(3 - 2\psi(\sigma_x, 0)\right)} F_{0x}(\sigma_x) \tag{42a}$$

$$F_{ay}(\sigma_y, \sigma_y) = \frac{3\rho\left(1 - \psi(\sigma_x, \sigma_y)\right)^2}{3\rho\left(1 - \psi(0, \sigma_y)\right)^2 + \psi(0, \sigma_y)\left(3 - 2\psi(0, \sigma_y)\right)} F_{0y}(\sigma_y) \tag{42b}$$

Proof. See Appendix A. □

Property 8 (Slip-velocity invariant sliding-forces)**.** *Combined-slip sliding forces and slip-velocity invariant pure-slip forces are related as:*

$$F_{sx}(\sigma_x, \sigma_y) = F_{0x}(\sigma_{0x}^{\text{vel}})\theta(\psi(\sigma_{0x}^{\text{vel}}, 0))|\cos(\beta')| \cdot \Gamma_x^{\text{vel}} \tag{43a}$$

$$F_{sy}(\sigma_x, \sigma_y) = F_{0y}(\sigma_{0y}^{\text{vel}})\theta(\psi(0, \sigma_{0y}^{\text{vel}}))|\sin(\beta')| \cdot \Gamma_y^{\text{vel}} \tag{43b}$$

with

$$\Gamma_x^{\text{vel}} \triangleq \begin{cases} \dfrac{\left(3 - 2\psi(\sigma_x, \sigma_y)\right)}{\left(3 - 2\psi(\sigma_{0x}^{\text{vel}}, 0)\right)} \left(|\cos(\beta)| \cos(\beta^\circ) + \dfrac{\sigma_x^\circ}{\sigma_y^\circ} |\sin(\beta)| \sin(\beta^\circ) \right)^2 \\ \cdot \left(\dfrac{1}{\rho_v} \sqrt{(1 + \sigma_x)^2 + \sigma_y^2} - \sqrt{\sigma_x^2 + \sigma_y^2}\, \text{sgn}(\sigma_x) \right)^2 \\ \qquad\qquad \textit{if } \psi(\sigma_x, \sigma_y) < 1,\ \psi(\sigma_{0x}^{\text{vel}}, 0) < 1 \\ \psi^{-2}(\sigma_{0x}^{\text{vel}}, 0)(3 - 2\psi(\sigma_{0x}^{\text{vel}}, 0))^{-1} \quad \textit{if } \psi(\sigma_x, \sigma_y) \geq 1,\ \psi(\sigma_{0x}^{\text{vel}}, 0) < 1 \\ \psi^2(\sigma_x, \sigma_y)(3 - 2\psi(\sigma_x, \sigma_y)) \qquad \textit{if } \psi(\sigma_x, \sigma_y) < 1,\ \psi(\sigma_{0x}^{\text{vel}}, 0) \geq 1 \\ 1 \qquad\qquad \textit{if } \psi(\sigma_x, \sigma_y) \geq 1,\ \psi(\sigma_{0x}^{\text{vel}}, 0) \geq 1 \end{cases} \tag{44a}$$

and

$$\Gamma_y^{\text{vel}} \triangleq \begin{cases} \dfrac{(3-2\psi(\sigma_x,\sigma_y))}{(3-2\psi(0,\sigma_{0y}^{\text{vel}}))}\left(\dfrac{\sigma_y^\circ}{\sigma_x^\circ}|\cos(\beta)|\cos(\beta^\circ)+|\sin(\beta)|\sin(\beta^\circ)\right)^2 \\ \cdot\left(\left(\dfrac{1}{\rho_v}\right)^2\left((1+\sigma_x)^2+\sigma_y^2\right)-(\sigma_x^2+\sigma_y^2)\right) \\ \qquad\qquad \text{if } \psi(\sigma_x,\sigma_y)<1,\ \psi(0,\sigma_{0y}^{\text{vel}})<1 \\ \psi^{-2}(0,\sigma_{0y}^{\text{vel}})(3-2\psi(\sigma_{0y}^{\text{vel}},0))^{-1} \quad \text{if } \psi(\sigma_x,\sigma_y)\geq 1,\ \psi(\sigma_{0y}^{\text{vel}},0)<1 \\ \psi^2(\sigma_x,\sigma_y)(3-2\psi(\sigma_x,\sigma_y)) \quad \text{if } \psi(\sigma_x,\sigma_y)<1,\ \psi(0,\sigma_{0y}^{\text{vel}})\geq 1 \\ 1 \quad \text{if } \psi(\sigma_x,\sigma_y)\geq 1,\ \psi(0,\sigma_{0y}^{\text{vel}})\geq 1 \end{cases} \tag{44b}$$

Proof. See Appendix A. □

It is necessary to formulate the scaling factors for different cases, to avoid numerically infeasible expressions at small slips. Note also, that to obtain the same slip velocity for combined and pure slips, it is necessary to take into account the corresponding wheel-travel velocities, v and v_0, which may be different.

Property 9 (Region-invariant adhesion forces)**.** *Combined-slip adhesion forces at partial sliding ($\psi(\sigma_x,\sigma_y)<1$) are related to region-invariant pure-slip forces as:*

$$F_{ax}(\sigma_x,\sigma_y) = \Big(1-\theta\big(\psi(\sigma_x,\sigma_y)\big)\Big)\,F_{0x}(\sigma_{0x}^{\text{reg}})\,\cos(\beta^\circ) \tag{45a}$$

$$F_{ay}(\sigma_y,\sigma_y) = \Big(1-\theta\big(\psi(\sigma_x,\sigma_y)\big)\Big)\,F_{0y}(\sigma_{0y}^{\text{reg}})\,\sin(\beta^\circ) \tag{45b}$$

with β° given by (16).

Proof. See Appendix A. □

Property 10 (Region-invariant sliding forces)**.** *Combined-slip sliding forces are related to region-invariant pure-slip forces as:*

$$F_{sx}(\sigma_x,\sigma_y) = \theta\big(\psi(\sigma_x,\sigma_y)\big)\,F_{0x}(\sigma_{0x}^{\text{reg}})\,|\cos(\beta')| \tag{46a}$$

$$F_{sy}(\sigma_x,\sigma_y) = \theta\big(\psi(\sigma_x,\sigma_y)\big)\,F_{0y}(\sigma_{0y}^{\text{reg}})\,|\sin(\beta')| \tag{46b}$$

Proof. See Appendix A. □

Property 11 (Region-invariant aligning torque)**.** *The combined-slip self-aligning torque is related to the region-invariant pure-slip lateral force and aligning moment as:*

$$M_z'(\sigma_x,\sigma_y) = M_{0z}(\sigma_{0y}^{\text{reg}})|\sin(\beta')| + \Big(1-\theta\big(\psi(0,\sigma_{0y}^{\text{reg}})\big)\Big)\, t_{0a}(\sigma_{0y}^{\text{reg}})F_{0y}(\sigma_{0y}^{\text{reg}})(\sin(\beta^\circ)-|\sin(\beta')|) \quad (47)$$

where

$$t_{0a}(\sigma_{0y}^{\text{reg}}) = \frac{C_z}{C_y}(4\psi(\sigma_x,\sigma_y)-1) \quad (48)$$

Proof. See Appendix A. □

Property 12 (Region-invariant additional aligning-torque)**.** *The combined-slip deformation torque is related to the region-invariant pure-slip lateral force and aligning moment as:*

$$M_{az}''(\sigma_x,\sigma_y) = \frac{12\,a\rho^2\,(1-\psi(\sigma_x,\sigma_y))^3}{\Gamma_{xy}^2}\left(\frac{1}{C_x}-\frac{1}{C_y}\right)F_{0x}(\sigma_{0x}^{\text{reg}})F_{0y}(\sigma_{0y}^{\text{reg}}) \quad (49a)$$

at partial sliding ($\psi(\sigma_x,\sigma_y)<1$), otherwise $M_{az}''(\sigma_x,\sigma_y)=0$ and

$$M_{sz}''(\sigma_x,\sigma_y) = \frac{6\Gamma_z}{5\Gamma_{xy}}a\left(\frac{1}{C_x}-\frac{1}{C_y}\right)F_{0x}(\sigma_{0x}^{\text{reg}})F_{0y}(\sigma_{0y}^{\text{reg}})|\cos(\beta')||\sin(\beta')| \quad (49b)$$

where

$$\Gamma_{xy} \triangleq 3\rho\,(1-\psi(\sigma_x,\sigma_y))^2 + \psi(\sigma_x,\sigma_y)\,(3-2\psi(\sigma_x,\sigma_y)) \quad (50a)$$

$$\Gamma_z \triangleq \big(10-15\psi(\sigma_x,\sigma_y)+6\psi^2(\sigma_x,\sigma_y)\big)\,\psi(\sigma_x,\sigma_y) \quad (50b)$$

at partial sliding ($\psi(\sigma_x,\sigma_y)<1$), otherwise $\Gamma_{zx}=\Gamma_{zy}=\Gamma_z=1$.

Proof. See Appendix A. □

4. Semi-Empirical Combined Slip Model

The proposed model is constructed by using the theoretical properties of Section 3, with $F_{0x}(\sigma_x)$, $F_{0y}(\sigma_y)$, and $M_{0z}(\sigma_y)$ replaced by empirical pure-slip models $\hat{F}_{0x}$, $\hat{F}_{0y}$, and $\hat{M}_{0z}$. In essence, this means that Property 3 is used to estimate sliding and adhesion force components from the empirical pure-slip data. Thereafter, combined forces are constructed using scaling factors based on the brush model.

Semi-empirical model for combined slip. *A semi-empirical combined-slip model based on pure-slip scaling is obtained by using empirical pure-slip models $\hat{F}_{0x}$, $\hat{F}_{0y}$, and $\hat{M}_{0z}$, in combination with:*

- *The deformation-invariant scaling of Property 7, to compute the semi-empirical adhesion force $\hat{F}_{ax}$,*
- *The slip-velocity invariant scaling of Property 8, to compute the velocity-dependent semi-empirical sliding force $\hat{F}_{sx}$,*
- *Property 11 together with 12, based on the region-invariant pure-slip, to compute the semi-empirical aligning-moments $\hat{M}'_z$ and $\hat{M}''_z$.*

The total semi-empirical combined-slip forces are then given by $\hat{F}_x = \hat{F}_{ax} + \hat{F}_{sx}$ and $\hat{F}_y = \hat{F}_{ay} + \hat{F}_{sy}$. The total semi-empirical combined-slip aligning-moment is given by $\hat{M}_z = \hat{M}'_z + \hat{M}''_z$.

The region-invariant pure slip is used for the aligning torque to keep the model simple. A torque model based on deformation-invariant and slip-velocity invariant slips is described in [Gäfvert and Svendenius, 2003]. An alternative model for the forces based on region-invariant scaling can also be stated. This model does not include any velocity dependence.

Semi-empirical model for combined slip (Alternative)**.** *An alternative semi-empirical combined-slip model based on pure-slip scaling is obtained by using empirical pure-slip models $\hat{F}_{0x}$, $\hat{F}_{0y}$, and $\hat{M}_{0z}$, in combination with:*

- *Property 9, to compute the semi-empirical adhesion force $\hat{F}_{ax}$,*
- *Property 10, to compute the semi-empirical sliding force $\hat{F}_{sx}$.*

The total semi-empirical combined-slip forces are then given by $\hat{F}_x = \hat{F}_{ax} + \hat{F}_{sx}$ and $\hat{F}_y = \hat{F}_{ay} + \hat{F}_{sy}$.

Empirical pure-slip models often include several effects that are not described by the brush-model. The most prominent are the normal-load dependence, effects due to carcass flexibility, and the apparent slip dependence, or in fact velocity dependence, of the sliding friction. Still, these

effects will be included in the forces from the combined-slip model, as they are, essentially, scalings of the empirical forces. In the special case of pure slips at $v = v_0$, the scaling factors are unity and these effects will be reproduced exactly.

Slip Parameterization

The brush model and, consequently, the proposed model are based on the slip $\bar{\sigma}$, defined by (5). This slip has the disadvantage of growing to infinity at wheel-lock. When implementing the proposed model a more practical slip has to be used. It is assumed that the empirical pure-slip models are parameterized in λ and α, defined by (6) and (1), as $\hat{F}_{0x}(\lambda)$, $\hat{F}_{0y}(\alpha)$ and $\hat{M}_{0z}(\alpha)$, which is common practice. It is also assumed that λ and α are used as inputs to the combined-slip model. Implementation of the model based on λ and α is described in Appendix B.

Parameters

Only four additional parameters are needed in the model, which all clear physical interpretations: σ_x°, σ_y°, ρ, and ρ_v. The parameters σ_x° and σ_y°, or correspondingly, λ° and α°, describe the pure slips where transition from partial to full sliding occur. They are needed to compute the normalized slip, $\psi(\sigma_x, \sigma_y)$. A common assumption is that these transitions occur when the tire forces obtain their maxima. Hence, the parameters may simply be set to the slip values corresponding to the maxima of $\hat{F}_{0x}$ and $\hat{F}_{0y}$. The parameter ρ reflects the ratio between the static and kinetic friction coefficients. It is not uncommon to assume these to be equal, i.e. $\rho = 1$, and this will normally give satisfactory results. For pure-slip models with very pronounced force peaks then $\rho > 1$ may be more accurate. The parameter ρ_v describes the ratio of the actual wheel-travel velocity, v, and the wheel-travel velocity for which the empirical pure-slip model is valid, v_0. If v_0 is not known, then setting $\rho_v = 1$ will neglect any velocity dependence. This is the common assumption in most other models.

Magic Formula

The proposed model is convenient to use with Magic Formula pure-slip models [Bakker *et al.*, 1987, Bakker *et al.*, 1989]. In this case $\hat{F}_{0x}(\lambda)$, $\hat{F}_{0y}(\alpha)$ and $\hat{M}_{0z}(\alpha)$ are described by the Magic Formula equations

$$f(\xi) = D \sin\Big(C \arctan\big((1-E)\xi + (E/B)\arctan(B\xi)\big)\Big) \tag{51}$$

where (ξ, f) is (λ, F_{0x}), (α, F_{0y}), or (α, M_{0z}), with the possibly normal-load dependent parameters $(B, C, D, E)_{x,y,z}$, correspondingly. Since the force-peak values are given by the Magic-Formula parameters as $\hat{F}_{0x}^* = (D)_x$,

$\hat{F}^*_{0y} = (D)_y$ and the cornering and braking stiffnesses as $\hat{C}_\lambda = (BCD)_x$, $\hat{C}_\alpha = (BCD)_y$, the parameters σ°_x and σ°_y may be conveniently estimated as

$$\hat{\sigma}^\circ_x = \frac{3}{(BC)_x} \tag{52a}$$

$$\hat{\sigma}^\circ_y = \frac{3 \cdot 180}{(BC)_y \pi} \tag{52b}$$

This follows from (14), using (25), under the assumption that $\rho = 1$, since the peak forces then are given by $F^*_{0x} = F_z \mu_{sx} = F_z \mu_{kx}$ and $F^*_{0y} = F_z \mu_{sy} = F_z \mu_{ky}$. The factor $180/\pi$ is due to α being expressed in degrees.

Compensation for Flexible Carcass. The brush model described in Section 2 is based on the assumption on a stiff carcass. In reality, the carcass is flexible and exhibits significant deformation laterally. One effect of the flexible carcass is that Equation (14), and therefore also (52b), no longer hold. It is, however, possible to adjust (52b) to compensate for the carcass flexibility.

Let $F'_y(\sigma_y)$ denote the lateral tire-force for a tire with flexible carcass. Regard, again, how the deformation $\delta_y(x)$ of an infinitesimal bristle element at position x in the adhesion region is described by (9). This is a purely kinematic relation that holds also in the case of a flexible carcass. However, in this case $\delta_y(x)$ is the sum of the bristle and carcass deflections $\delta_{yb}(x)$ and $\delta_{yc}(x)$:

$$\delta_y(x) = \delta_{yb}(x) + \delta_{yc}(x) \tag{53}$$

There are several theories on how to describe the carcass deformation and a simple linear approach is used here:

$$\delta_{yc}(x) = -\frac{F'_y(\sigma_y)}{C_c}(a - x) \tag{54}$$

where C_c is a carcass stiffness. From (53), (54) and (9) it holds that

$$\delta_{yb}(x) = -\left(\sigma_y - \frac{F'_y(\sigma_y)}{C_c}\right)(a - x) \tag{55}$$

The force acting on the bristle element at x is then

$$dF'_y(x) = c_{py}\delta_{yb}(x)\,dx = -c_{py}\left(\sigma_y - \frac{F'_y(\sigma_y)}{C_c}\right)(a - x)\,dx \tag{56}$$

At the break-away point, $x = x_s$, the static friction constraint (12) at pure lateral slip, together with the the pressure distribution (13), gives

$$c_{py}\left(\sigma_y - \frac{F_y'(\sigma_y)}{C_c}\right) = \frac{3\mu_{\mathrm{s}y}F_z}{4a^3}\left(a + x_s\right) \tag{57}$$

When the entire contact patch slides, then $x_s = a$, $\sigma_y = \sigma_y^{\circ\prime}$ and $F_y'(\sigma_y^{\circ\prime}) = \mu_{\mathrm{k}y}F_z$. Solving (57) for $\sigma_y^{\circ\prime}$ under these conditions gives

$$\sigma_y^{\circ\prime} = \frac{3F_z\mu_{\mathrm{s}y}}{2a^2c_{py}} + \frac{\mu_{\mathrm{k}y}F_z}{C_c} = F_z\left(\frac{3\mu_{\mathrm{s}y}}{2a^2c_{py}} + \frac{\mu_{\mathrm{k}y}}{C_c}\right) \tag{58}$$

The next step is to derive an expression for the relation between C_c and the cornering stiffness C_y'. At very small slips, $\sigma_y \approx 0$, there is no sliding in the contact patch and the tire force only consists of adhesion force. Therefore,

$$\begin{aligned} C_y' = \left.\frac{dF_y'(\sigma_y)}{d\sigma_y}\right|_{\sigma_y=0} &= -\left.\frac{d}{d\sigma_y}\left(\int_{-a}^{a} c_{py}\delta_b(x)\,dx\right)\right|_{\sigma_y=0} \\ &= 2a^2c_{py}\left(1 - \frac{1}{C_c}\left.\frac{dF_y'(\sigma_y)}{d\sigma_y}\right|_{\sigma_y=0}\right) = C_y\left(1 - \frac{C_y'}{C_c}\right) \end{aligned} \tag{59}$$

where $2a^2c_{py} = C_y$, according to (25). Hence,

$$C_y' = \frac{C_y}{1 + \dfrac{C_y}{C_c}} \tag{60}$$

Further, the rubber is assumed to be isotropic, $c_{py} = c_{px}$, and hence, $C_y = C_x$, which is a realistic assumption. The carcass stiffness can be calculated from (60) as

$$C_c = \frac{C_xC_y'}{C_x - C_y'} \tag{61}$$

Using (61), the limit-slip adjusted for carcass deformation (58) can be written as

$$\sigma_y^{\circ\prime} = F_z\mu_{\mathrm{s}y}\left(\frac{2}{C_x} + \frac{1}{C_y'}\right) \tag{62}$$

Regard that $F_y^* = F_z\mu_{\mathrm{s}y}$. The Magic Formula parameters give $\hat{F}_y^* = D_y$, $\hat{C}_x = (BCD)_x$, and $\hat{C}_y' = (BCD)_y \cdot 180/\pi$. Finally,

$$\hat{\sigma}_y^{\circ\prime} = D_y\left(\frac{2}{(BCD)_x} + \frac{\pi}{180\cdot(BCD)_y}\right) \tag{63}$$

Utilizing Braking Data to Generate Driving Data

In the proposed model it is assumed that the given longitudinal pure-slip model, $\hat{F}_{0x}(\lambda)$, is valid both at braking ($\lambda > 0$) and driving ($\lambda < 0$). The Magic Formula is an odd function and therefore $\hat{F}_{0x}(\lambda) = -\hat{F}_{0x}(-\lambda)$. Generally, this is not in accordance with empirical observations. If the pure-slip model is assumed to be valid for braking, then the following procedure suggests how to modify the argument to the pure-slip model at driving, so that a more accurate force is obtained.

The brush model states that the adhesion force developed at braking with a slip $\sigma_x > 0$ will have the same size and the opposite sign, as the force generated at driving with slip $-\sigma_x$:

$$F_{0ax}(\sigma_x) = -F_{0ax}(-\sigma_x) \tag{64}$$

With $\sigma_x = -\lambda_{\text{driving}}/(1-\lambda_{\text{driving}})$ and $\sigma_x = \lambda_{\text{braking}}/(1-\lambda_{\text{braking}})$ this means that

$$\lambda_{\text{braking}} = -\frac{\lambda_{\text{driving}}}{1-2\lambda_{\text{driving}}} \tag{65}$$

Hence, when computing adhesion forces using an empirical model for braking, the pure-slip forces

$$\begin{cases} -\hat{F}_{0x}\left(\dfrac{-\lambda}{1-2\lambda}\right) & \lambda < 0 \\ \hat{F}_{0x}(\lambda) & \lambda \geq 0 \end{cases} \tag{66}$$

are best used for driving and braking, respectively. For the sliding case it is more natural to let the force depend on the relative velocity $v_s = \lambda v_x$. Then, $-v_s$ will simply correspond to $-\lambda$. Therefore, when computing sliding forces using an empirical model for braking, the pure-slip forces

$$\begin{cases} -\hat{F}_{0x}(-\lambda) & \lambda < 0 \\ \hat{F}_{0x}(\lambda) & \lambda \geq 0 \end{cases} \tag{67}$$

are best used for driving and braking respectively. The effect of the described procedure is visible for negative λ in the upper plot of Figure 6.

5. Results and Discussion

Empirical data from a truck-tire is used to exemplify the proposed model. The data is provided by Volvo Truck Corporation [Edlund, 1991]. It consists of pure-slip data with corresponding Magic Formula parameters, and two sets of combined-slip data for fixed slip-angles, $\alpha = 4.7$ deg and $\alpha = 9.8$ deg, with varying slip ratios, λ, in the range 0–100%. Unless stated otherwise, the proposed model is used with default parameters.

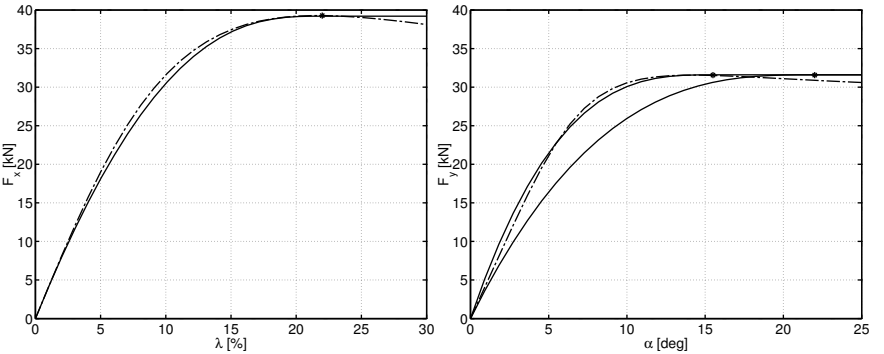

Figure 5 Brush-model forces (solid) and Magic-Formula empirical data (dash-dotted).The limit slips are marked with asterisks. Left: Longitudinal force with σ_x° given by (52a). Right: Lateral force with σ_y° given by (52b) (lower solid line) or (63) (upper solid line). Note that the adjustment for flexible carcass results in significantly better fit.

The Brush Model

The agreement between the theoretical brush-model and empirical pure-slip data is shown in Figure 5. The effect of adjusting the lateral limit-slip to compensate for flexible carcass is illustrated. Note that the original expression for σ_y°, (52b), overestimates the limit slip, resulting in poor fit with empirical data. With the modified expression, (63), the fit is equally good as in the longitudinal direction. The agreement supports the use of the brush model as the basis for the combined-slip model.

Validation

The combined-slip model reduces exactly to the empirical model at pure slip when $\rho_v = 1$, see Figure 6. In Figure 7, the proposed model and the alternative model are compared with the combined-slip empirical data. The resulting tire-forces for constant $\alpha = 4.7$ deg and $\alpha = 9.8$ deg, as λ is swept from 0 to 100%, are plotted. Both models show equally good agreement with the validation data. Differences are best explained by the limitations of the empirical data, as noted below. Empirical data for the combined-slip self-aligning moment were not available.

Measurement Data

The combined-slip validation data consists of two measurement sets which exhibit a few obvious discrepancies. The pure-slip lateral forces in the combined-slip data sets do not fully agree with the corresponding forces in the pure-slip data sets. This may be seen by comparing the data points

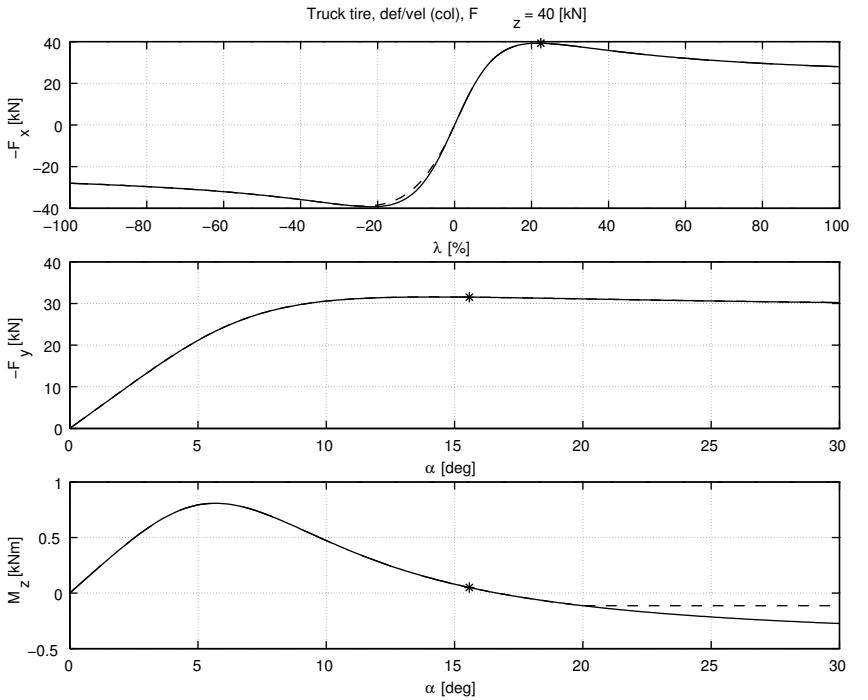

Figure 6 Pure-slip forces and self-aligning moment for the empirical pure-slip model (solid lines) and the proposed semi-empirical model (dashed lines). The limit slips λ° and α° are marked by asterisks. The deviation at $\lambda < 0$ in the longitudinal direction results from using the procedure of Section 4 at driving. Since the Magic Formula is valid for $|\alpha| < 20$ deg, the aligning-torque data for $\alpha = \pm 20$ deg are used for larger slip-angles.

for F_y at $F_x = 0$ in Figure 7, with the corresponding points at $\alpha = 4.7$ deg and $\alpha = 9.8$ deg in Figure 6. The data points corresponding to large slips, $\lambda \approx 100\%$ (wheel-lock), in the combined-slip data sets seem strange and the corresponding directions of the tire force are difficult to explain. At $\alpha = 4.7$ deg, the direction of the force at $\lambda = 100\%$ is 1.28 deg, where it should be 4.7 deg for isotropic kinetic friction. The corresponding direction for $\alpha = 9.8$ deg is 5.1 deg. No obvious explanation has been found to this and it is assumed that it is the result of measurement errors.*

Measurement data for the lateral force and the self-aligning moment is only available for $|\alpha| < 20$ deg and the validity of the fitted Magic

*Models based on assumptions of anisotropic friction and the Maximum Dissipation-Rate principle were investigated in [Gäfvert and Svendenius, 2003], without yielding better results.

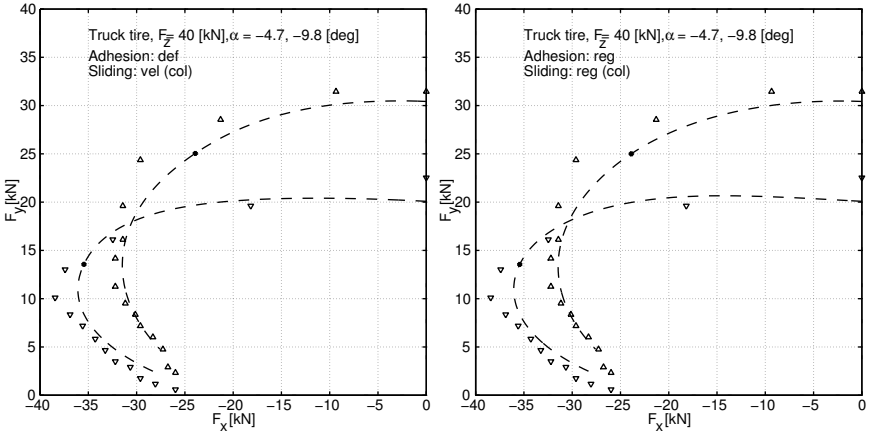

Figure 7 Comparison between the proposed combined-slip model (left), the alternative model (right), and combined-slip measurements for $\alpha = 4.7$ deg and $\alpha = 9.8$ deg, as λ is swept from 0 to 100%. The results are very close.

Formula model is, therefore, limited to this range. For $|\alpha| > 20$ deg the data for $\alpha = \pm 20$ deg are used for the self-aligning torque.

Combined-Slip Examples

In Figure 8, the resulting forces and aligning moments are shown for fixed slip-angles and varying slip-ratio, ranging from -100% (driving) to 100% (braking with locked wheels). The adhesion and sliding contributions are shown separately for the forces. The adhesion force is dominating at small slip magnitudes and vanishes at the point of full sliding. For the aligning-moment, the contributions from $M'(\sigma_x, \sigma_y)$ and $M''(\sigma_x, \sigma_y)$ are shown separately. It is clear the main contribution to the self-aligning torque is given by M'_z, which is the part resulting from the non-symmetric distribution of the lateral force. The additional part M''_z, resulting from tire deformation, is smaller. This part is, however, probably underestimated, since only the deformation resulting from bristle deflection is accounted for. Note the asymmetrical characteristics with respect to driving ($\lambda < 0$) and braking ($\lambda > 0$). This is, essentially, an effect of the relation between the slips $\bar{\sigma}$ and $(\lambda, \tan(\alpha))$, see (7). The combined-slip forces and moments agree qualitatively well with observations reported in e.g. [Pacejka, 2002]. Figure 9 shows the corresponding case with fixed slip ratio, as the slip angle is swept from 0 to 30 deg. Negative slip angles are not shown since the characteristics are symmetrical. Figure 10 is similar to Figure 7 and shows the combined forces at constant slip angles, as the slip ratio is swept from -100% to 100%. For small slips, the direction of

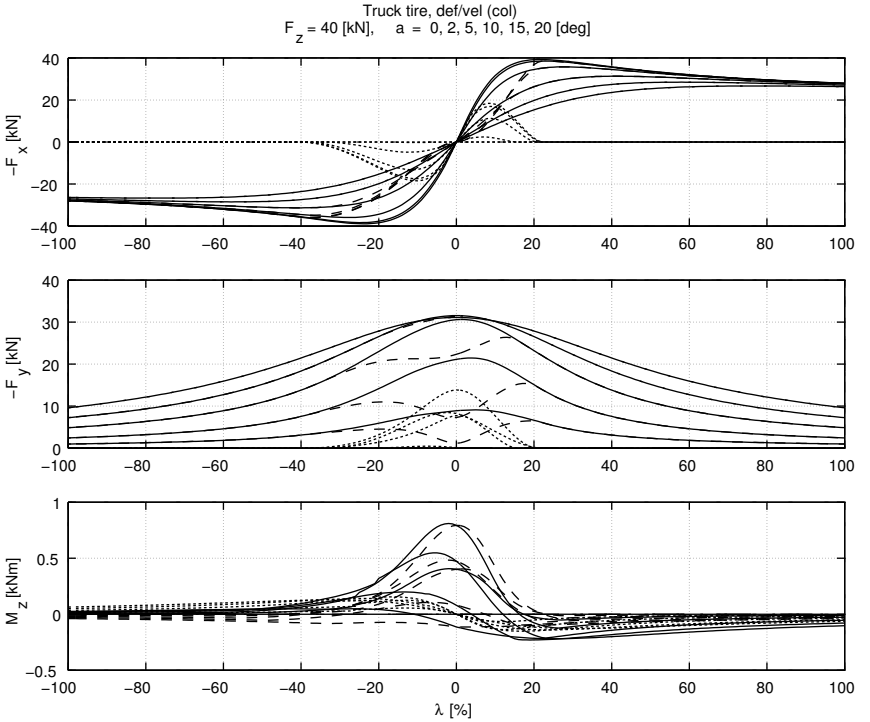

Figure 8 Combined-slip forces for fixed α, as λ is swept from -100% (driving) to 100% (braking with locked wheels). Top: F_x (solid), F_{ax} (dotted), F_{sx} (dashed). Middle: F_y (solid), F_{ay} (dotted), F_{sy} (dashed). Bottom: M_z (solid), M_z' (dashed), M_z'' (dotted).

the tire-force is, primarily, determined by the stiffness characteristics of the tire. For larger slips, when the sliding friction dominates, the force is collinear with the slip vector. This gradual change in orientation of the force with increasing slips is an important feature of the model.

Velocity Dependence

The dependence on wheel-travel velocity is illustrated in Figures 11 and 12, by varying the parameter $\rho_v = v/v_0$. The pure-slip Magic-Formula model is calibrated with data from tire measurements at the wheel-travel velocity v_0.[†] Tire forces are shown for velocities that are 1–4 times the wheel-travel velocity of the pure-slip model. The results agree qualita-

[†]The exact value of v_0 is not known for this data. Measurements are normally performed at around 10 m/s on the used test-bed.

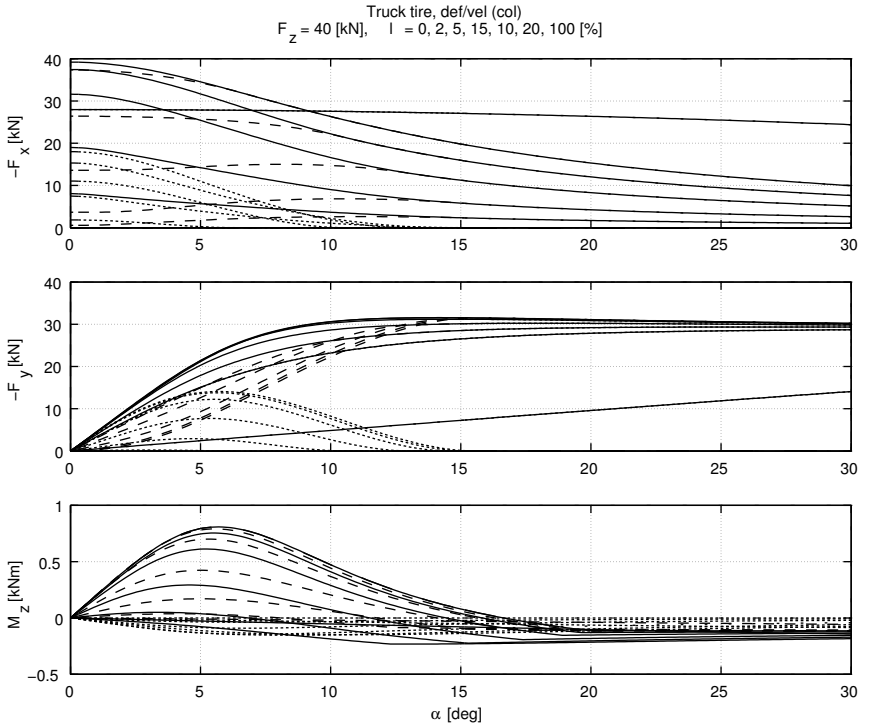

Figure 9 Combined-slip forces for fixed λ, as α is swept from 0 to 30 deg. Top: F_x (solid), F_{ax} (dotted), F_{sx} (dashed). Middle: F_y (solid), F_{ay} (dotted), F_{sy} (dashed). Bottom: M_z (solid), M_z' (dashed), M_z'' (dotted).

tively well with what is reported in e.g. [Pacejka, 2002]. The slip velocity at a combined slip λ, α, is $v_s = v\left(\left(\lambda\cos(\alpha)\right)^2 + \sin^2(\alpha)\right)^{1/2}$. Corresponding pure slips for the empirical pure-slip model are given by $v_s = \lambda_0^{\text{vel}} v_0$ and $v_s = \sin(\alpha_0^{\text{vel}}) v_0$. These expressions can be solved for pure slips $\lambda_0^{\text{vel}} < 100\%$ and $\alpha_0^{\text{vel}} < 90$ deg, when $\rho_v^2\left(\lambda^2\cos^2(\alpha) + \sin^2(\alpha)\right) \leq 1$, see (B.4). If $\rho_v^2\left(\lambda^2\cos^2(\alpha) + \sin^2(\alpha)\right) > 1$, which may occur only if $\rho_v > 1$ or α_0^{vel} is outside the valid range of the pure-slip model, then extrapolation is necessary. A straightforward method is to use the end points of the models. The Magic Formula with sound parameters may be used with slip ratios $\lambda_0^{\text{vel}} > 100\%$.

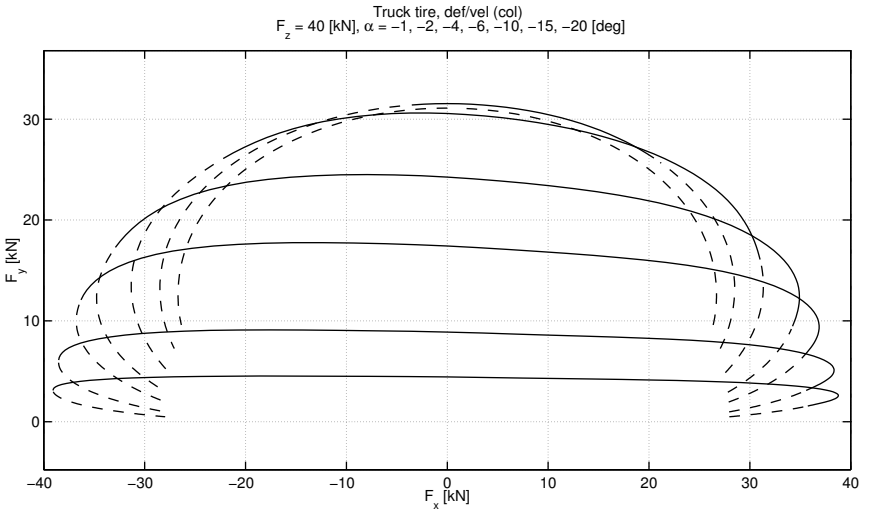

Figure 10 Combined-slip forces for fixed α as λ is swept from -100% (driving) to 100% (braking with locked wheels). Partial sliding (solid); Full sliding (dashed).

Normal-Force Dependence

Normal-force dependence is straightforwardly included in the proposed model by using normal-force dependent pure-slip models, such as the standard Magic Formula.

Relations to Other Models

In Figure 13, comparisons are shown between the proposed model, the BPL model [Bakker *et al.*, 1989], and the COMBINATOR model [Schuring *et al.*, 1996]. It is clear that the proposed model performs better than the COMBINATOR model at small slip-angles. The COMBINATOR model may be written as

$$F_x = F\cos(\beta) \tag{68a}$$

$$F_y = F\sin(\beta) \tag{68b}$$

where

$$F = F_{0x}(s)\cos^2(\beta) + F_{0y}(s)\sin^2(\beta) \tag{69}$$

The resulting force is always collinear with the slip vector, which is an assumption with weak physical motivation. In the COMBINATOR model the lateral force initially increases, as a longitudinal slip is applied. This is a result of the assumption of a collinear combined-slip tire-force in the full

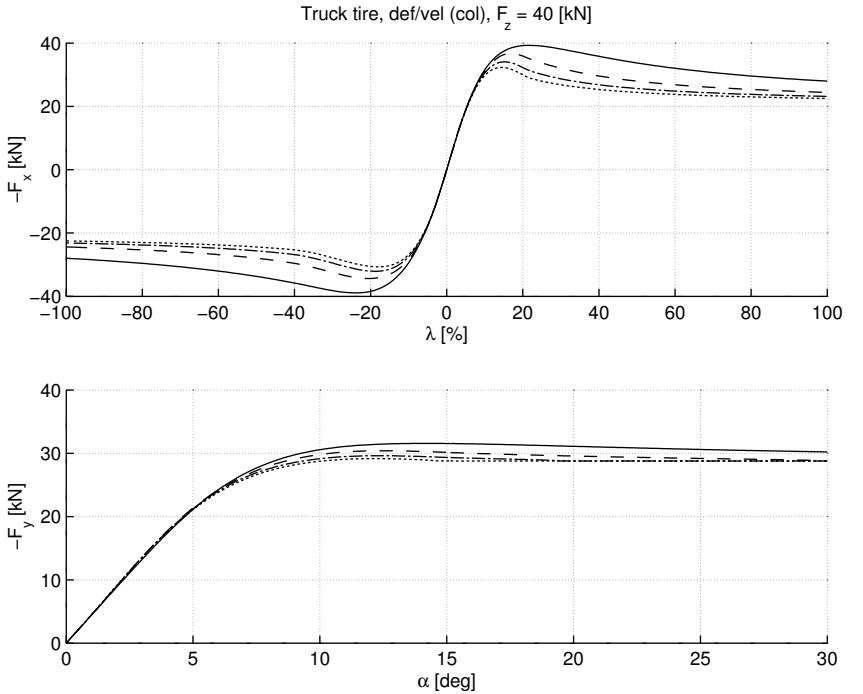

Figure 11 Velocity dependence at pure slips. Velocities: $v = v_0$ (solid), $2v_0$ (dashed), $3v_0$ (dash-dotted), $4v_0$ (dotted).

slip range, in combination with the use of the combined-slip magnitude in the empirical pure-slip models.

It is clear from the figure that the proposed region-invariant model is very similar to the BPL model. One major reason for this is that also the BPL model is based on the region-invariant pure slips. The BPL equations

$$F_x = \cos\left((1-\vartheta)\beta^\circ + \vartheta\beta\right) F'_{0x} \tag{70a}$$

$$F_y = \sin\left((1-\vartheta)\beta^\circ + \vartheta\beta\right) F'_{0y} \tag{70b}$$

where

$$\vartheta \triangleq \frac{2}{\pi}\arctan(q_1\psi^2(\sigma_x,\sigma_y))) \tag{71}$$

with the parameter q_1 and

$$F'_{0x} \triangleq F_{0x}(\sigma_{0x}^{\text{reg}}) - \text{sat}(\psi(\sigma_x,\sigma_y))(F_{0x}(\sigma_{0x}^{\text{reg}}) - F_{0y}(\sigma_{0y}^{\text{reg}}))\sin^2(\beta^\circ) \tag{72a}$$

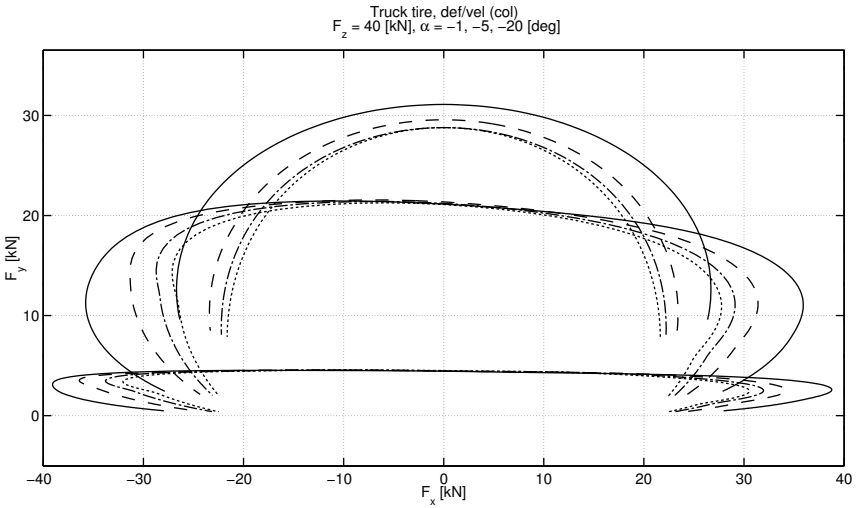

Figure 12 Velocity dependence at combined slips. Velocities: $v = v_0$ (solid), $2v_0$ (dashed), $3v_0$ (dash-dotted), $4v_0$ (dotted).

$$F'_{0y} \triangleq F_{0y}(\sigma_{0y}^{\text{reg}}) + \text{sat}(\psi(\sigma_x, \sigma_y))(F_{0x}(\sigma_{0x}^{\text{reg}}) - F_{0y}(\sigma_{0y}^{\text{reg}}))\cos^2(\beta^\circ) \quad (72b)$$

are also clearly similar to the region-invariant model, which may be written as

$$F_x = ((1-\theta)\cos(\beta^\circ) + \theta\cos(\beta'))\, F_{0x} \quad (73a)$$

$$F_y = ((1-\theta)\sin(\beta^\circ) + \theta\sin(\beta'))\, F_{0y} \quad (73b)$$

with θ given by (39). In the BPL model, the orientation of the force is obtained by a convex combination of angles, corresponding to adhesion and sliding, respectively. In the region-invariant model the direction is obtained by a convex combination of adhesion and sliding forces. The resultant magnitude for the BPL model at full sliding is

$$F = F_{0x}\cos^2(\beta^\circ) + F_{0y}\sin^2(\beta^\circ) \quad (74)$$

Note that β° depends on the tire stiffnesses and the static friction coefficient, which do not have any physical significance at full sliding. For the region-invariant model the corresponding magnitude is

$$F_{0x}F_{0y}/\sqrt{F_{0x}^2\sin^2(\beta) + F_{0y}^2\cos^2(\beta)} \quad (75)$$

Compare with (69) for the COMBINATOR model.

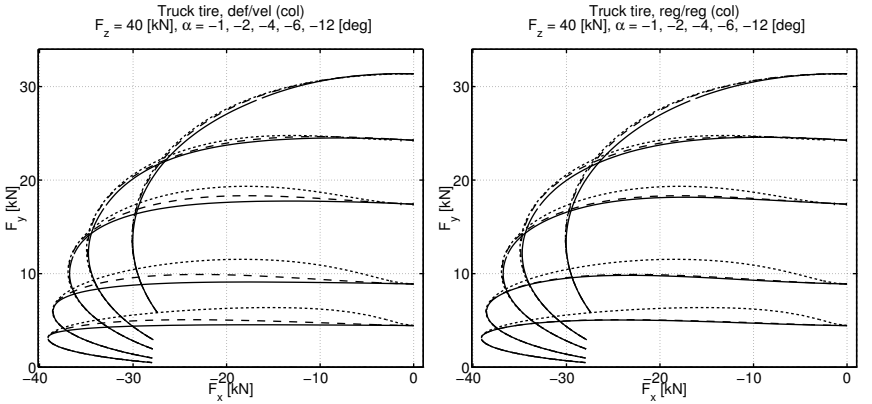

Figure 13 Combined-slip forces with the proposed model (left plot, solid line) and the alternative model (right plot, solid line) for fixed slip angles α and varying λ ranging from 0 to 100% (braking with locked wheels). The COMBINATOR model (dotted line) and the BPL model (dashed line) are shown for comparison.

6. Summary

The proposed model is based on understanding of the physical sources of tire forces, as given by the theoretical rigid-carcass brush-model. Based on the brush model, the combined-slip forces may be described with a scaling of corresponding empirical pure-slip forces. A procedure is presented to separate the adhesive and sliding components of empirical pure-slip forces. From these, combined-slip forces for adhesion and sliding are constructed separately. This makes sense, since these forces are the result of different physical processes. The pure slips corresponding to a particular combined slip are, therefore, also chosen differently for adhesion and sliding. For adhesion forces, the pure slips which describe the same tread-deformation as the combined slip are used. For sliding forces, the pure slips which result in the same slip velocity as the combined slip are used. This approach includes velocity dependence in the model. An alternative model is also suggested, that is based on pure slips corresponding to the same size of adhesion and sliding regions, as the combined slip. Velocity dependence is then not included, but simpler expressions are obtained. The alternative model is closely related to the previously published BPL model [Bakker *et al.*, 1989]. In [Gäfvert and Svendenius, 2003] several models based on other choices of pure slips are investigated.

The brush model includes the major physical processes which are responsible for the build-up of tire-forces. Additional effects resulting from, e.g., varying normal-load, flexible carcass, and velocity-dependent sliding-

forces are implicitly included in the proposed model, if they are present in the empirical pure-slip models.

Four parameters, each with clear physical interpretations, are introduced in the model. Two of these parameters describe when transition to full sliding occurs and are of central significance. They can be computed automatically from the pure-slip models. One parameter describes the relation between the actual wheel-travel velocity and the wheel-travel velocity of the pure-slip model. The default value of unity means that velocity dependence is neglected. Finally, one parameter describes an assumption on the relation between static and kinetic friction. This parameter is set to unity as default and describes a freedom to adjust the model response, if desired. The Magic Formula is a natural choice of empirical pure-slip model and it integrates well with the proposed model.

Good results are obtained in validation with empirical data. Relations to similar, previously published combined-slip models are analyzed. It is concluded that the suggested model is more accurate than the COMBINATOR model [Schuring *et al.*, 1996]. The BPL model [Bakker *et al.*, 1989] presented together with the Magic Formula show similar agreement, but does not include velocity dependence.

Preliminary studies on extensions of the model to include effects of a flexible carcass are ongoing. The goal is to include simple models of dynamic effects and transient behavior, resulting from carcass deformation. It is believed that this is possible without introducing overwhelming complexity. Effects of ply-steer, conicity, and camber have not been regarded to this point. It is expected that also these may be included in the future.

Acknowledgements

The first author would like to thank the Swedish Research Council for funding this work under contract 2001-3020. The second author would like to thank Haldex Brake Products AB, Landskrona, Sweden, for funding and support.

References

Bakker, E., L. Nyborg, and H. B. Pacejka (1987): “Tyre modelling for use in vehicle dynamics studies.” SAE Techical Paper 870421.

Bakker, E., H. B. Pacejka, and L. Lidner (1989): “A new tire model with an application in vehicle dynamics studies.” SAE Technical Paper 890087.

Edlund, S. (1991): "Tyre models: Subreport -91." Technical Report. Volvo Truck Corporation. Confidential.

Gäfvert, M. and J. Svendenius (2003): "Construction of semi-empirical tire models for combined slip." Technical Report ISRN LUTFD2/TFRT--7606--SE. Department of Automatic Control, Lund Institute of Technology, Sweden.

Gim, G. and P. E. Nikravesh (1991): "An analytical model of pneumatic tyres for vehicle dynamics simulations. part 2: Comprehensive slips." *International Journal of Vehicle Design*, **12:1**, pp. 19–39.

ISO 8855 (1991): "Road vehicles — vehicle dynamics and road-holding ability — vocabulary."

Pacejka, H. B. (1988): "Modelling of the pneumatic tyre and its impact on vehicle dynamic behavior." Technical Report i72B. Technische Universiteit Delft.

Pacejka, H. B. (2002): *Tyre and Vehicle Dynamics*. Butterworth-Heinemann.

SAE Recommended Practice J670e (1976): "Vehicle dynamics terminology."

Schuring, D. J., W. Pelz, and M. G. Pottinger (1996): "A model for combined tire cornering and braking forces." In *Investigations and Analysis in Vehicle Dynamics and Simulation*, pp. 61–83. SAE International. SAE Technical Paper 960180.

Wong, J. Y. (2001): *Theory of Ground Vehicles*, 3 edition. John Wiley & Sons.

A. Proofs

Proof of Property 3. Regard the longitudinal sliding force at partial sliding ($\psi(\sigma_x, 0) < 1$):

$$
\begin{aligned}
F_{0sx}(\sigma_x) &= \frac{F_{0sx}(\sigma_x)}{F_{0ax}(\sigma_x) + F_{0sx}(\sigma_x)} F_{0x}(\sigma_x) \\
&= \frac{\mu_{\mathrm{kx}} F_z \psi^2(\sigma_x, 0)\,(3 - 2\psi(\sigma_x, 0))\,\mathrm{sgn}(\sigma_x)}{2a^2 c_{px(\sigma_x)} \sigma_x\,(1 - \psi(\sigma_x, 0))^2 + \mu_{\mathrm{kx}} F_z \psi_x^2(\sigma_x, 0)\,(3 - 2\psi(\sigma_x, 0))\,\mathrm{sgn}(\sigma_x)} \\
&\quad \cdot F_{0x}(\sigma_x) = \frac{\psi(\sigma_x, 0)\,(3 - 2\psi(\sigma_x, 0))}{3\rho\,(1 - \psi(\sigma_x, 0))^2 + \psi(\sigma_x, 0)\,(3 - 2\psi(\sigma_x, 0))} F_{0x}(\sigma_x) \\
&= \theta(\psi(\sigma_x, 0)) F_{0x}(\sigma_x) \qquad \text{(A.1)}
\end{aligned}
$$

The second equality is from insertion of (18), (23) and (20). The third equality is obtained by using (14). At full sliding ($\psi(\sigma_x, 0) \geq 1$), then $F_{0ax}(\sigma_x) = 0$ and $F_{0sx}(\sigma_x) = F_{0x}(\sigma_x)$. Hence, $\theta(\psi(\sigma_x, 0)) = 1$. The expressions for $\theta(\psi(0, \sigma_y))$ are derived analogously. □

Proof of Property 4. Follows from (9), which expresses that the deformations in the longitudinal and lateral directions are independent. □

Proof of Property 5. The relative velocity v_s, at the combined slip (σ_x, σ_y) and the wheel-travel velocity v, is given by

$$v_s = \sqrt{v_{sx}^2 + v_{sy}^2} = \frac{v\sqrt{\sigma_x^2 + \sigma_y^2}}{\sqrt{(1+\sigma_x)^2 + \sigma_y^2}} \tag{A.2}$$

For the pure longitudinal slip at the wheel-travel velocity $v_{0x} = v_0$ and the slip velocity $v_{0sx} = v_s \operatorname{sgn}(\sigma_x)$, the slip definition gives:

$$\sigma_{0x}^{\text{vel}} = \frac{v_{0sx}}{v_{0x} - v_{0sx}} = \frac{v_s \operatorname{sgn}(\sigma_x)}{v_0 - v_s \operatorname{sgn}(\sigma_x)} = \frac{\sqrt{\sigma_x^2 + \sigma_y^2}\operatorname{sgn}(\sigma_x)}{\frac{v_0}{v}\sqrt{(1+\sigma_x)^2 + \sigma_y^2} - \sqrt{\sigma_x^2 + \sigma_y^2}\operatorname{sgn}(\sigma_x)} \tag{A.3}$$

The lateral slip is shown analogously. □

Proof of Property 6. It follows from (41) and (15) that $\psi(\sigma_{0x}^{\text{reg}}, 0) = \psi(\sigma_x, \sigma_y)$. From (17) it then follows that $x_s(\sigma_{0x}^{\text{reg}}, 0) = x_s(\sigma_x, \sigma_y)$. The lateral slip is shown correspondingly. □

Proof of Property 7. Assume $\sigma_x \neq 0$, then $F_{0ax}(\sigma_x) \neq 0$, since $\psi(\sigma_x, \sigma_y) < 1$, and

$$\begin{aligned} F_{ax}(\sigma_x, \sigma_y) = F_{0ax}(\sigma_x)\frac{F_{ax}(\sigma_x, \sigma_y)}{F_{0ax}(\sigma_x)} &= F_{0ax}(\sigma_x)\left(\frac{1-\psi(\sigma_x, \sigma_y)}{1-\psi(\sigma_x, 0)}\right)^2 \\ &= (1 - \theta(\psi(\sigma_x, 0)))\, F_{0x}(\sigma_x)\left(\frac{1-\psi(\sigma_x, \sigma_y)}{1-\psi(\sigma_x, 0)}\right)^2 \end{aligned} \tag{A.4}$$

The second equality is due to (18) and (19). Inserting (39) and rearranging yields (42). The scaling factor is non-singular at $\sigma_x = 0$. If $\sigma_x = 0$ then $F_{0ax}(\sigma_x) = F_{ax}(\sigma_x, \sigma_y) = 0$, which is fulfilled by (42). The lateral expression is derived analogously. □

Proof of Property 8. Regard the longitudinal force and assume $\sigma_x \neq 0$. Then using (23), (20), and (21):

$$F_{sx}(\sigma_x,\sigma_y) \quad = \quad F_{0sx}(\sigma_{0x}^{\mathrm{vel}})\frac{F_{sx}(\sigma_x,\sigma_y)}{F_{0sx}(\sigma_{0x}^{\mathrm{vel}})} \quad = \quad |\cos(\beta')|F_{0x}(\sigma_{0x}^{\mathrm{vel}})\theta_x(\sigma_{0x}^{\mathrm{vel}})\Gamma_x^{\mathrm{vel}} \tag{A.5}$$

with

$$\Gamma_x^{\mathrm{vel}} = \frac{F_{sz}(\sigma_x,\sigma_y)}{F_{sz}(\sigma_{0x}^{\mathrm{vel}},0)} = \frac{\psi^2(\sigma_x,\sigma_y)(3-2\psi(\sigma_x,\sigma_y))}{\psi^2(\sigma_{0x}^{\mathrm{vel}},0)(3-2\psi(\sigma_{0x}^{\mathrm{vel}},0))} \tag{A.6}$$

when $\psi(\sigma_x,\sigma_y) < 1$ and $\psi(\sigma_{0x}^{\mathrm{vel}},0) < 1$. If $\psi(\sigma_x,\sigma_y) \geq 1$ then $\psi(\sigma_x,\sigma_y)$ is replaced by unity in (A.6). If $\psi(\sigma_{0x}^{\mathrm{vel}},0) \geq 1$ then $\psi(\sigma_{0x}^{\mathrm{vel}},0)$ replaced by unity. Rewriting is necessary to avoid singular numerics at small slips, due to the factor $\psi^2(\sigma_x^{\mathrm{vel}},0)$ in the denominator. When $\psi(\sigma_x,\sigma_y) < 1$ and $\psi(\sigma_{0x}^{\mathrm{vel}},0) < 1$ then using (15) and (40):

$$\frac{\psi^2(\sigma_x,\sigma_y)}{\psi^2(\sigma_{0x}^{\mathrm{vel}},0)} = \left(\frac{\cos(\beta)}{\cos(\beta^\circ)}\right)^2\left(\frac{v_0}{v}\sqrt{(1+\sigma_x)^2+\sigma_y^2} - \mathrm{sgn}(\sigma_x)\sqrt{\sigma_x^2+\sigma_y^2}\right)^2 \tag{A.7}$$

Using the definitions of β and β°, it is then easily verified that:

$$\frac{\cos(\beta)}{\cos(\beta^\circ)} = \sigma_x^\circ\left(\frac{1}{\sigma_x^\circ}\cos(\beta)\cos(\beta^\circ) + \frac{1}{\sigma_y^\circ}\sin(\beta)\sin(\beta^\circ)\right) \tag{A.8}$$

From the definition it is clear that $\psi(\sigma_x^{\mathrm{vel}},0)$ is not close to zero when $\psi(\sigma_x,\sigma_y) \geq 1$. For $\sigma_x = 0$, then $\cos(\beta) = 0$, which give $F_{0sx} = 0$. The lateral expressions are shown analogously. □

Proof of Property 9. Assume $\sigma_x \neq 0$ and $\psi(\sigma_x,\sigma_y) < 1$, then:

$$\begin{aligned} F_{ax}(\sigma_x,\sigma_y) &= F_{0ax}(\sigma_{0x}^{\mathrm{reg}})\frac{F_{ax}(\sigma_x,\sigma_y)}{F_{0ax}(\sigma_{0x}^{\mathrm{reg}})} = F_{0ax}(\sigma_{0x}^{\mathrm{reg}})\frac{\sigma_x\left(1-\psi(\sigma_x,\sigma_y)\right)}{\sigma_{0x}^{\mathrm{reg}}\left(1-\psi(\sigma_{0x}^{\mathrm{reg}},0)\right)} \\ &= F_{0ax}(\sigma_{0x}^{\mathrm{reg}})\frac{\sigma_x\left(1-\psi(\sigma_x,\sigma_y)\right)}{\sigma_x^\circ\psi(\sigma_x,\sigma_y)\left(1-\psi(\sigma_x,\sigma_y)\right)} = F_{0ax}(\sigma_{0x}^{\mathrm{reg}})\frac{\psi(\sigma_x,0)}{\psi(\sigma_x,\sigma_y)} \\ &= (1-\theta(\psi(\sigma_{0x}^{\mathrm{reg}},0)))F_{0x}(\sigma_{0x}^{\mathrm{reg}})\cos(\beta^\circ) \end{aligned} \tag{A.9}$$

Finally, recall that $\psi(\sigma_{0x}^{\mathrm{reg}},0) = \psi(\sigma_x,\sigma_y)$. If $\sigma_x = 0$ then $F_{0ax}(\sigma_x) = F_{ax}(\sigma_x,\sigma_y) = 0$, which is fulfilled by (45) since then $\sin(\beta^\circ) = 0$. The lateral expression is derived analogously. □

Proof of Property 10. Regard the longitudinal force. Then, using (23), (20), (41), and (21):

$$\begin{aligned} F_{sx}(\sigma_x,\sigma_y) &= -cos(\beta)\mu_{\mathrm{kx}}F_{sz}(\sigma_x,\sigma_y) = -cos(\beta)\mu_{\mathrm{kx}}F_{sz}(\sigma_{0x}^{\mathrm{reg}},0) \\ &= F_{0sx}(\sigma_{0x}^{\mathrm{reg}})\cos(\beta) = \theta(\psi(\sigma_{0x}^{\mathrm{reg}},0))F_{0x}(\sigma_{0x}^{\mathrm{reg}})\cos(\beta) \end{aligned} \tag{A.10a}$$

Again, recall that $\psi(\sigma_{0x}^{\mathrm{reg}},0) = \psi(\sigma_x,\sigma_y)$. The lateral expression is derived analogously. □

Proof of Property 11. From (35) and Property 6 it is clear that the adhesion and sliding pneumatic trails for the combined slip (σ_x, σ_y) is equal to the pneumatic trails for the pure slip $(0, \sigma_{0y}^{\text{reg}})$: $t'_a(\sigma_x, \sigma_y) = t'_{0a}(\sigma_{0y}^{\text{reg}})$, $t'_s(\sigma_x, \sigma_y) = t'_{0s}(\sigma_{0y}^{\text{reg}})$. By using (29) and (35):

$$\begin{aligned}
M'_z(\sigma_x, \sigma_y) &= t'_s(\sigma_x, \sigma_y) F_{sy}(\sigma_x, \sigma_y) + t'_a(\sigma_x, \sigma_y) F_{ay}(\sigma_x, \sigma_y) \\
&= t'_{0s}(\sigma_{0y}^{\text{reg}}) F_{0sy}(\sigma_{0y}^{\text{reg}}) \frac{F_{sy}(\sigma_x, \sigma_y)}{F_{0sy}(\sigma_{0y}^{\text{reg}})} + t'_{0a}(\sigma_{0y}^{\text{reg}}) F_{0ay}(\sigma_{0y}^{\text{reg}}) \frac{F_{ay}(\sigma_x, \sigma_y)}{F_{0ay}(\sigma_{0y}^{\text{reg}})} \\
&= M_{0sz}(\sigma_{0y}^{\text{reg}}) \sin(\beta') + t'_{0a}(\sigma_{0y}^{\text{reg}}) F_{0ay}(\sigma_{0y}^{\text{reg}}) \sin(\beta^\circ) \\
&= (M_{0z}(\sigma_{0y}^{\text{reg}}) - t'_{0a}(\sigma_{0y}^{\text{reg}}) F_{0ay}(\sigma_{0y}^{\text{reg}})) \sin(\beta') + t'_{0a}(\sigma_{0y}^{\text{reg}}) F_{0ay}(\sigma_{0y}^{\text{reg}}) \sin(\beta^\circ) \\
&= M_{0z}(\sigma_{0y}^{\text{reg}}) \sin(\beta') + t'_{0a}(\sigma_{0y}^{\text{reg}}) (1 - \theta(\psi(0, \sigma_{0y}^{\text{reg}}))) F_{0y}(\sigma_{0y}^{\text{reg}}) (\sin(\beta^\circ) - \sin(\beta'))
\end{aligned} \tag{A.11}$$

where Properties 9 and 10 have been used in the third equality. The contact length a in (35) may be computed from (36) as $a = 3C_z/C_y$. □

Proof of Property 12. For the adhesion part M''_{az} of the deformation torque the expressions for the adhesion forces given by Property 9 are inserted in (31). The factor $(1-\psi(\sigma_x, \sigma_y))$, present in the denominator of (31) and in the numerator of θ, can then be cancelled to avoid numerical problems. The sliding part (M''_{sz}) is given by (32) and the sliding forces by Property 10. Again, the factor $\psi(\sigma_x, \sigma_y)$ present in the denominator of (32) and in the numerator of $1 - \theta$ can be canceled. □

B. Implementation Details

This appendix is intended as guidance when implementing the suggested model. Pseudo-code implementations of Properties 7–12 are given, together with the formulation of some model expressions in λ and α.

Pseudo-Code Implementations

Pseudo-code implementation of the semi-empirical deformation-invariant adhesion-force model based on Property 7.

Require: λ, α, $(B, C, D, E)_{x,y}$, ρ
Produces: $\hat{F}_{ax}(\lambda, \alpha)$, $\hat{F}_{ay}(\lambda, \alpha)$
  if $\psi(\sigma_x, \sigma_y) < 1$, (B.7) **then**
    σ_x, $\sigma_y \Leftarrow \lambda$, α, (B.1)
    σ_x°, $\sigma_y^\circ \Leftarrow (B, C, D)_{x,y}$, (52a), (63)
    $\psi(\sigma_x, \sigma_y) \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, (15)
    λ_0, $\alpha_0 \Leftarrow \lambda$, α, (B.2)
    $\hat{F}_{0x}(\lambda_0)$, $\hat{F}_{0y}(\alpha_0) \Leftarrow \lambda_0$, α_0, $(B, C, D, E)_{x,y}$, (51), (66)
    σ_{0x}, $\sigma_{0y} \Leftarrow \sigma_x$, σ_y, Property 4
    $\psi(\sigma_{0x}, 0)$, $\psi(0, \sigma_{0y}) \Leftarrow \sigma_{0x}$, σ_{0y}, σ_x°, σ_y°, (15)

$\hat{F}_{ax}(\lambda,\alpha)$, $\hat{F}_{ay}(\lambda,\alpha) \Leftarrow \hat{F}_{0x}(\lambda_0)$, $\hat{F}_{0y}(\alpha_0)$, $\psi(\sigma_x,\sigma_y)$, $\psi(\sigma_{0x},0)$, $\psi(0,\sigma_{0y})$, ρ, (42)
else
$\hat{F}_{ax}(\lambda,\alpha) := 0$ and $\hat{F}_{ay}(\lambda,\alpha) := 0$
end if

Pseudo-code implementation of the semi-empirical slip-velocity invariant slide-force model based on Property 8.

Require: λ, α, $(B,C,D,E)_{x,y}$, ρ, ρ_v
Produces: $\hat{F}_{sx}(\lambda,\alpha)$, $\hat{F}_{sy}(\lambda,\alpha)$
λ_0^{vel}, $\alpha_0^{\mathrm{vel}} \Leftarrow \lambda$, α, (B.4)
$\hat{F}_{0x}(\lambda_0^{\mathrm{vel}})$, $\hat{F}_{0y}(\alpha_0^{\mathrm{vel}}) \Leftarrow \lambda_0^{\mathrm{vel}}$, α_0^{vel}, $(B,C,D,E)_{x,y}$, (51)
σ_x°, $\sigma_y^\circ \Leftarrow (B,C,D)_{x,y}$, (52a), (63)
$\beta \Leftarrow \lambda$, α, (4)
if $\psi(\sigma_x,\sigma_y) < 1$ (B.7) **then**
σ_x, $\sigma_y \Leftarrow \lambda$, α, (B.1)
$\psi(\sigma_x,\sigma_y)$, $\psi(\sigma_x,0)$, $\psi(0,\sigma_y) \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, (15)
$\beta^\circ \Leftarrow \psi(\sigma_x,0)$, $\psi(0,\sigma_y)$, (16)
end if
if $\psi(\sigma_{0x}^{\mathrm{vel}},0) < 1$, (B.8) **then**
$\sigma_{0x}^{\mathrm{vel}} \Leftarrow \alpha$, λ, (B.3a)
$\psi(\sigma_{0x}^{\mathrm{vel}},0) \Leftarrow \sigma_{0x}^{\mathrm{vel}}$ (15)
$\theta(\psi(\sigma_{0x}^{\mathrm{vel}},0)) \Leftarrow \psi(\sigma_{0x}^{\mathrm{vel}},0)$, ρ, (39)
else
$\theta(\psi(\sigma_{0x}^{\mathrm{vel}},0)) := 1$
end if
if $\psi(0,\sigma_{0y}^{\mathrm{vel}}) < 1$, (B.8) **then**
$\sigma_{0y}^{\mathrm{vel}} \Leftarrow \alpha$, λ, (B.3b)
$\psi(0,\sigma_{0y}^{\mathrm{vel}}) \Leftarrow \sigma_{0y}^{\mathrm{vel}}$, (15)
$\theta(\psi(0,\sigma_{0y}^{\mathrm{vel}})) \Leftarrow \sigma_{0y}^{\mathrm{vel}}$, ρ, (39)
else
$\theta_y(\psi(0,\sigma_{0y}^{\mathrm{vel}})) := 1$
end if
Γ_x^{vel}, $\Gamma_y^{\mathrm{vel}} \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, β, β°, $\psi(\sigma_x,\sigma_y)$, $\psi(\sigma_{0x}^{\mathrm{vel}},0)$, $\psi(0,\sigma_{0y}^{\mathrm{vel}})$, ρ_v, (44)
$\hat{F}_{0sx}(\lambda_0^{\mathrm{vel}}) := \theta(\psi(\sigma_{0x}^{\mathrm{vel}},0))\hat{F}_{0x}(\lambda_0^{\mathrm{vel}})\Gamma_x^{\mathrm{vel}}$
$\hat{F}_{0sy}(\alpha_0^{\mathrm{vel}}) := \theta(\psi(0,\sigma_{0y}^{\mathrm{vel}}))\hat{F}_{0y}(\alpha_0^{\mathrm{vel}})\Gamma_y^{\mathrm{vel}}$
$\beta' \Leftarrow \lambda$, α, $\hat{F}_{0sx}(\lambda_0^{\mathrm{vel}})$, $\hat{F}_{0sy}(\alpha_0^{\mathrm{vel}})$, (22) using $\mu_{\mathrm{kx}}/\mu_{\mathrm{ky}} = \hat{F}_{0sx}(\lambda_0^{\mathrm{vel}})/\hat{F}_{0sy}(\alpha_0^{\mathrm{vel}})$ and $v_{sy}/v_{sx} = \sin(\alpha)/(\lambda\cos(\alpha))$.
$\hat{F}_{sx}(\lambda,\alpha)$, $\hat{F}_{sy}(\lambda,\alpha) \Leftarrow \hat{F}_{0x}(\lambda_0^{\mathrm{vel}})$, $\hat{F}_{0y}(\alpha_0^{\mathrm{vel}})$, $\theta_x(\sigma_{0x}^{\mathrm{reg}})$, $\theta_y(\sigma_{0y}^{\mathrm{reg}})$, β', (43)

Pseudo-code implementation of the semi-empirical region-invariant adhesion-force model based on Property 9.

Require: λ, α, $(B,C,D,E)_{x,y}$, ρ
Produces: $\hat{F}_{ax}(\lambda,\alpha)$, $\hat{F}_{ay}(\lambda,\alpha)$
if $\psi(\sigma_x,\sigma_y) < 1$, (B.7) **then**
λ_0^{reg}, $\alpha_0^{\mathrm{reg}} \Leftarrow \lambda$, α, (B.6)
$\hat{F}_{0x}(\lambda_0^{\mathrm{reg}})$, $\hat{F}_{0y}(\alpha_0^{\mathrm{reg}}) \Leftarrow \lambda_0^{\mathrm{reg}}$, α_0^{reg}, $(B,C,D,E)_{x,y}$, (51), (66)

σ_x°, $\sigma_y^\circ \Leftarrow (B,C,D)_{x,y}$, (52a), (63)
σ_x, $\sigma_y \Leftarrow \lambda$, α, (B.1)
$\psi(\sigma_x,\sigma_y)$, $\psi(\sigma_x,0)$, $\psi(0,\sigma_y) \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, (15)
$\beta^\circ \Leftarrow \psi(\sigma_x,0)$, $\psi(0,\sigma_y)$, (16)
$\theta(\psi(\sigma_x,\sigma_y)) \Leftarrow \psi(\sigma_x,\sigma_y)$, ρ, (39)
$\hat{F}_{ax}(\lambda,\alpha)$, $\hat{F}_{ay}(\lambda,\alpha) \Leftarrow \hat{F}_{0x}(\lambda_0^{\text{reg}})$, $\hat{F}_{0y}(\alpha_0^{\text{reg}})$, β°, $\theta(\psi(\sigma_x,\sigma_y))$, (45)
else
$\hat{F}_{ax}(\lambda,\alpha) := 0$ and $\hat{F}_{ay}(\lambda,\alpha) := 0$
end if

Pseudo-code implementation of the semi-empirical region-invariant sliding-force model based on Property 10.

Require: λ, α, $(B,C,D,E)_{x,y}$, ρ
Produces: $\hat{F}_{sx}(\lambda,\alpha)$, $\hat{F}_{sy}(\lambda,\alpha)$
λ_0^{reg}, $\alpha_0^{\text{reg}} \Leftarrow \lambda$, α, (B.6)
$\hat{F}_{0x}(\lambda_0^{\text{reg}})$, $\hat{F}_{0y}(\alpha_0^{\text{reg}}) \Leftarrow \lambda_0^{\text{reg}}$, α_0^{reg}, $(B,C,D,E)_{x,y}$, (51)
σ_x°, $\sigma_y^\circ \Leftarrow (B,C,D)_{x,y}$, (52a), (63)
if $\psi(\sigma_x,\sigma_y) < 1$ (B.7) **then**
σ_x, $\sigma_y \Leftarrow \lambda$, α, (B.1)
$\psi(\sigma_x,\sigma_y) \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, (15)
$\theta(\psi(\sigma_x,\sigma_y)) \Leftarrow \psi(\sigma_x,\sigma_y)$, ρ, (39)
else
$\theta(\psi(\sigma_x,\sigma_y)) := 1$
end if
$\hat{F}_{0sx}(\lambda_0^{\text{reg}}) := \theta(\psi(\sigma_x,\sigma_y))\hat{F}_{0x}(\lambda_0^{\text{reg}})$
$\hat{F}_{0sy}(\alpha_0^{\text{reg}}) := \theta(\psi(\sigma_x,\sigma_y))\hat{F}_{0y}(\alpha_0^{\text{reg}})$
$\beta' \Leftarrow \lambda$, α, $\hat{F}_{0sx}(\lambda_0^{\text{reg}})$, $\hat{F}_{0sy}(\alpha_0^{\text{reg}})$, (22) Using $\mu_{\text{kx}}/\mu_{\text{ky}} = \hat{F}_{0sx}(\lambda_0^{\text{reg}})/\hat{F}_{0sy}(\alpha_0^{\text{reg}})$ and $v_{sy}/v_{sx} = \sin(\alpha)/(\lambda\cos(\alpha))$
$\hat{F}_{sx}(\lambda,\alpha)$, $\hat{F}_{sy}(\lambda,\alpha) \Leftarrow \hat{F}_{0x}(\lambda_0^{\text{reg}})$, $\hat{F}_{0y}(\alpha_0^{\text{reg}})$, $\theta(\psi(\sigma_x,\sigma_y))$, β', (46)

Pseudo-code implementation of the semi-empirical region-invariant aligning-moment model based on Property 11.

Require: λ, α, $(B,C,D,E)_{y,z}$, ρ
Produces: $\hat{M}_z'(\lambda,\alpha)$
$\alpha_0^{\text{reg}} \Leftarrow \lambda$, α, (B.6)
$\hat{F}_{0y}(\alpha_0^{\text{reg}})$, $\hat{M}_{0z}(\alpha_0^{\text{reg}}) \Leftarrow \alpha_0^{\text{reg}}$, $(B,C,D,E)_{y,z}$, (51)
σ_x°, $\sigma_y^\circ \Leftarrow (B,C,D)_{x,y}$, (52a), (63)
$\beta \Leftarrow \lambda$, α, (4)
if $\psi(\sigma_x,\sigma_y) < 1$ (B.7) **then**
σ_x, $\sigma_y \Leftarrow \lambda$, α, (B.1)
$\psi(\sigma_x,\sigma_y)$, $\psi(\sigma_x,0)$, $\psi(0,\sigma_y) \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, (15)
$\beta^\circ \Leftarrow \psi(\sigma_x,0)$, $\psi(0,\sigma_y)$, (16)
$\theta(\psi(\sigma_x,\sigma_y)) \Leftarrow \psi(\sigma_x,\sigma_y)$, ρ, (39)
$t_{0a}(\sigma_{0y}^{\text{reg}}) \Leftarrow \psi(\sigma_x,\sigma_y)$, $(B,C,D)_y$, $(B,C,D)_z$, (48)
$\hat{M}_z'(\lambda,\alpha) \Leftarrow \sigma_{0y}^{\text{reg}}$, β, β°, $\hat{F}_{0y}(\alpha_0^{\text{reg}})\,\theta(\psi(\sigma_x,\sigma_y))$, $t_{0a}(\sigma_{0y}^{\text{reg}})$, (47)
else

$\hat{M}'_z(\lambda,\alpha) := \hat{M}_{0z}(\alpha_0^{\text{reg}})\sin(\beta)$, (47)
end if

Pseudo-code implementation of the semi-empirical region-invariant additional aligning-moment model based on Property 12.

Require: λ, α, $(B,C,D,E)_{x,y}$, ρ
Produces: $\hat{M}''_z(\lambda,\alpha)$
λ_0^{reg}, $\alpha_0^{\text{reg}} \Leftarrow \lambda$, α, (B.6)
$\hat{F}_{0x}(\lambda_0^{\text{reg}})$, $\hat{F}_{0y}(\alpha_0^{\text{reg}}) \Leftarrow \lambda_0^{\text{reg}}$, α_0^{reg}, $(B,C,D,E)_{x,y}$, (51)
σ_x°, $\sigma_y^\circ \Leftarrow (B,C,D)_{x,y}$, (52a), (63)
if $\psi(\sigma_x,\sigma_y) < 1$ (B.7) **then**
σ_x, $\sigma_y \Leftarrow \lambda$, α, (B.1)
$\psi(\sigma_x,\sigma_y)$, $\psi(\sigma_x,0)$, $\psi(0,\sigma_y) \Leftarrow \sigma_x$, σ_y, σ_x°, σ_y°, (15)
σ_{0x}^{reg}, $\sigma_{0y}^{\text{reg}} \Leftarrow \psi(\sigma_x,\sigma_y)$, σ_x°, σ_y°, (41)
$\theta(\psi(\sigma_x,\sigma_y)) \Leftarrow \psi(\sigma_x,\sigma_y)$, ρ, (39)
Γ_{xy}, Γ_z, $\Leftarrow \psi(\sigma_x,\sigma_y)$, ρ, (50)
$M''_{az}(\lambda,\alpha) \Leftarrow \psi(\sigma_x,\sigma_y)$, $\hat{F}_{0x}(\lambda_0^{\text{reg}})$, $\hat{F}_{0y}(\alpha_0^{\text{reg}})$, (49)
else
$\theta(\psi(\sigma_x,\sigma_y)) := 1$
$\Gamma_{xy} := 1$, $\Gamma_z := 1$
$M''_{az}(\lambda,\alpha) := 0$
end if
$\hat{F}_{0sx}(\lambda_0^{\text{reg}}) := \theta(\psi(\sigma_x,\sigma_y))\hat{F}_{0x}(\lambda_0^{\text{reg}})$
$\hat{F}_{0sy}(\alpha_0^{\text{reg}}) := \theta(\psi(\sigma_x,\sigma_y))\hat{F}_{0y}(\alpha_0^{\text{reg}})$
$\beta' \Leftarrow \lambda$, α, $\hat{F}_{0sx}(\lambda_0^{\text{reg}})$, $\hat{F}_{0sy}(\alpha_0^{\text{reg}})$, (22) using $\mu_{\text{kx}}/\mu_{\text{ky}} = \hat{F}_{0sx}(\lambda_0^{\text{reg}})/\hat{F}_{0sy}(\alpha_0^{\text{reg}})$ and $v_{sy}/v_{sx} = \sin(\alpha)/(\lambda\cos(\alpha))$.
$M''_{sz}(\lambda,\alpha) \Leftarrow \hat{F}_{0x}(\lambda_0^{\text{reg}})$, $\hat{F}_{0y}(\alpha_0^{\text{reg}})$, β', Γ_{xy}, Γ_z, (49)
$\hat{M}''_z(\lambda,\alpha) := \hat{M}''_{az}(\lambda,\alpha) + \hat{M}''_{sz}(\lambda,\alpha)$

Reformulations in λ and α

The following reformulations of some of the model expressions are needed for implementation. They are based on straightforward application of the slip transformations (7).

$$\sigma_x = \frac{\lambda}{1-\lambda} \tag{B.1a}$$

$$\sigma_y = \frac{\tan(\alpha)}{1-\lambda} \tag{B.1b}$$

$$\lambda_0 = \frac{\sigma_{0x}}{1+\sigma_{0x}} = \lambda \tag{B.2a}$$

$$\tan(\alpha_0) = \sigma_{0y} = \frac{\sin(\alpha)}{(1-\lambda)\cos(\alpha)} \tag{B.2b}$$

$$\sigma_{0x}^{\text{vel}} = \frac{\rho_v \sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)}\,\text{sgn}(\lambda)}{1 - \rho_v \sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)}\,\text{sgn}(\lambda)} \tag{B.3a}$$

$$\sigma_{0y}^{\text{vel}} = \frac{\rho_v \sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)}}{\sqrt{1 - \rho_v^2 \left((\lambda \cos(\alpha))^2 + \sin^2(\alpha)\right)}}\,\text{sgn}(\alpha) \tag{B.3b}$$

$$\lambda_0^{\text{vel}} = \rho_v \sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)}\,\text{sgn}(\lambda) \tag{B.4a}$$

$$\sin(\alpha_0^{\text{vel}}) = \rho_v \sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)}\,\text{sgn}(\alpha) \tag{B.4b}$$

$$\sigma_{0x}^{\text{reg}} = \sigma_x^\circ \sqrt{\left(\frac{\lambda}{1-\lambda}\frac{1}{\sigma_x^\circ}\right)^2 + \left(\frac{\tan(\alpha)}{1-\lambda}\frac{1}{\sigma_y^\circ}\right)^2}\,\text{sgn}(\lambda) \tag{B.5a}$$

$$\sigma_{0y}^{\text{reg}} = \sigma_y^\circ \sqrt{\left(\frac{\lambda}{1-\lambda}\frac{1}{\sigma_x^\circ}\right)^2 + \left(\frac{\tan(\alpha)}{1-\lambda}\frac{1}{\sigma_y^\circ}\right)^2}\,\text{sgn}(\alpha) \tag{B.5b}$$

$$\lambda_0^{\text{reg}} = \frac{\sigma_{0x}^{\text{reg}}}{1 + \sigma_{0x}^{\text{reg}}} = \frac{\sigma_x^\circ \sqrt{\left(\frac{\lambda\cos(\alpha)}{\sigma_x^\circ}\right)^2 + \left(\frac{\sin(\alpha)}{\sigma_y^\circ}\right)^2}\,\text{sgn}(\lambda)}{(1-\lambda)\,|\cos(\alpha)| + \sigma_x^\circ \sqrt{\left(\frac{\lambda\cos(\alpha)}{\sigma_x^\circ}\right)^2 + \left(\frac{\sin(\alpha)}{\sigma_y^\circ}\right)^2}\,\text{sgn}(\lambda)} \tag{B.6a}$$

$$\tan(\alpha_0^{\text{reg}}) = \sigma_{0y}^{\text{reg}} = \frac{\sigma_y^\circ \sqrt{\left(\frac{\lambda\cos(\alpha)}{\sigma_x^\circ}\right)^2 + \left(\frac{\sin(\alpha)}{\sigma_y^\circ}\right)^2}}{(1-\lambda)|\cos(\alpha)|}\,\text{sgn}(\alpha) \tag{B.6b}$$

$$\psi(\sigma_x, \sigma_y) < 1 \iff \left(\frac{\lambda}{\sigma_x^\circ}\cos(\alpha)\right)^2 + \left(\frac{\sin(\alpha)}{\sigma_y^\circ}\right)^2 < (1-\lambda)^2 \cos^2(\alpha) \tag{B.7}$$

$$\psi(\sigma_{0x}^{\text{vel}}, 0) < 1 \iff$$
$$v\sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)} < \sigma_x^\circ |v_0 - v\sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)}\,\text{sgn}(\lambda)| \tag{B.8a}$$

$$\psi(0, \sigma_{0y}^{\text{vel}}) < 1 \iff$$
$$v\sqrt{(\lambda \cos(\alpha))^2 + \sin^2(\alpha)} < \sigma_y^\circ \sqrt{v_0^2 - v^2\left((\lambda \cos(\alpha))^2 + \sin^2(\alpha)\right)} \tag{B.8b}$$

Paper III

Control of GDI Engines using Torque Feedback Exemplified by Simulations

Magnus Gäfvert Karl-Erik Årzén
Lars Malcolm Pedersen Bo Bernhardsson

Abstract

A novel approach to the control of a gasoline direct injection (GDI) engine is presented. The controller consists of a combination of sub-controllers, where torque feedback is a central part. The sub-controllers are with a few exceptions designed using simple linear feedback and feedforward control-design methods, in contrast to traditional table-based engine control. A switching strategy which maintains driving comfort during combustion-mode changes is also proposed. A novel silent extremum-controller is presented, and used to minimize the fuel consumption in stratified combustion mode. The controller has been evaluated with good results on the European driving cycle using a dynamic simulation model.

1. Introduction

Increased environmental requirements and demands on decreased fuel consumption have during recent years led to the development of new types of combustion engines within the car industry. One of these engine types is the gasoline direct injection (GDI) engine, also known as the DISI (Direct Injection, Spark Ignition) engine. The main difference between a GDI engine and a conventional port fuel injection (PFI) engine concerns where air and fuel are mixed, and when the mixing takes place. In a PFI engine the fuel is mixed with the air in the intake ports outside the cylinder. The air-fuel mixture is entered into the cylinders during the intake stroke. This gives rise to a homogeneous air-fuel mixture in the cylinders that, to minimize emissions, should be kept at stoichiometric ratio. In a GDI engine, instead, the fuel is injected directly into the cylinder late in the compression stroke. Direct injection in the compression stroke is combined with a stratified charge, that only needs to be sufficiently rich locally, close to the spark. This enables the use of extremely lean air-fuel mixtures which reduces fuel consumption. The difference between PFI and GDI engines is shown in Figure 1. GDI engines put high demands on the

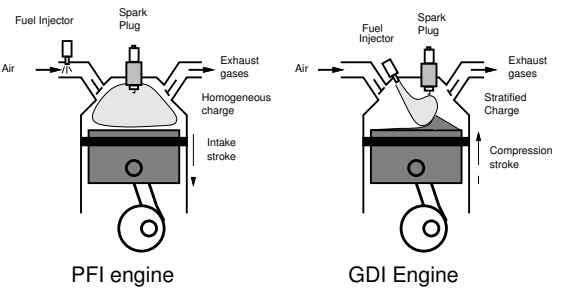

Figure 1 Differences between a PFI and a GDI engine. In a GDI engine the top of the piston head is shaped to create a fuel swirl that gives maximum fuel concentration at the spark.

timing of the injection. Good performance requires that exactly the right amount of fuel is injected at the correct time.

The stratified combustion mode in GDI engines can only be used at low or moderate engine loads and for low or moderate engine speeds, e.g., in low-speed urban traffic. For high loads the amount of fuel that needs to be injected is too large to be timely injected. For high engine speeds the time window when the fuel must be injected is too small. Therefore, GDI engines also use an alternative combustion mode, the homogeneous mode. Here, the fuel is still injected directly into the cylinder, but now

it is done in the intake stroke rather than in the compression stroke. In homogeneous mode, the GDI engine behaves similarly to a PFI engine.

GDI engines were commercially pioneered by Mitsubishi. However, currently most major car manufacturers have GDI engines in their production, e.g., Ford, GM, and VW. More information on the operation of GDI engines can be found in [Zao *et al.*, 1997].

Modern internal combustion engines constitute very complex systems in need of precise and robust control. The control objectives in engine control fall into two categories: tracking of reference values such as engine speed, engine torque or air-fuel ratio, and optimizing operating conditions such that fuel consumption and emissions are kept at a minimum. The GDI engine is more complex than an ordinary PFI engine and therefore requires a more advanced control system. Conventional control of car engines is normally based on extensive use of engine maps, i.e., multidimensional look-up tables that have been derived through extensive engine test-bench experiments [Gerhardt *et al.*, 1998]. This is an open-loop approach which can be sensitive to engine-to-engine variations and variations due to aging and wear. In this paper it is shown, using simulations based on a calibrated GDI engine model, that it is possible to instead employ simple feedback control strategies with a low number of tuning parameters and yet achieve good control performance, provided the effective torque is available for feedback.

For optimal fuel economy it is essential to run the GDI engine in stratified combustion mode whenever possible. Smooth mode-switching between the stratified and homogeneous combustion modes is important for driver comfort. In this paper a mode-switching strategy is presented that minimizes torque bumps and reduces fuel consumption.

In stratified mode the air-fuel ratio may be chosen within an interval. The air-fuel ratio can then be chosen in an optimal manner with respect to engine efficiency and emissions. In this paper a novel silent extremum-controller is presented, which is used to find the air-fuel ratio that generates a desired torque with the smallest possible fuel consumption. Engine control is a well known application area for extremum control, see e.g. [Draper and Li, 1951, Kiencke, 1987]. A problem with the lean-burn combustion is the increased formation of nitrogen oxides. To reduce the pollutant emissions exhaust gas recirculation (EGR) is normally introduced in the engine. Then there is need for optimizing the amount of recirculated gas with respect to engine efficiency and emissions. The extremum-control strategy is also extended to the case where EGR is present.

The work presented in this paper has been performed within the EU/Esprit Long Term Research project FAMIMO. Within FAMIMO a Matlab/Simulink™ dynamic simulation model of a 2 liters GDI engine

provided by Siemens Automotive has served as a common benchmark for evaluation of different control strategies. This paper is based on the conference papers [Gäfvert *et al.*, 2000b, Gäfvert *et al.*, 2000a] and the book chapter [Johansson and Rantzer, 2002]. Other related results are presented in [Passaquay, 2000] and [Mollov *et al.*, 2001].

The paper is organized as follows. The overall structure of the benchmark used in the simulation studies is described in Section 2. The GDI engine model used as the basis for the control design is overviewed in Section 3. Controller designs and the overall structure of the control system are presented in Section 4. The results of the controller on simulations of European Driving Cycle scenarios are presented in Section 5. Comparisons are made with other control designs. Finally, conclusions are given in Section 6.

2. The benchmark

The overall structure of the GDI benchmark is shown in Figure 2. The

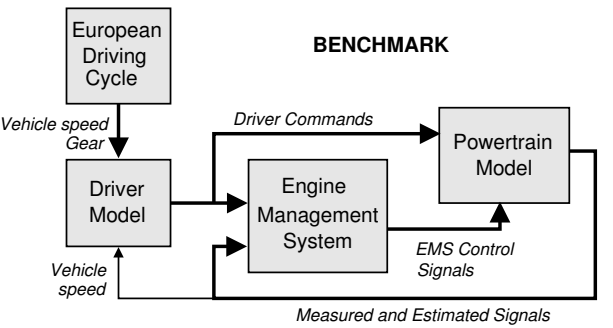

Figure 2 Structure of the GDI benchmark.

goal is to design an Engine Management System (EMS), or controller, that follows the command signals from the driver while minimizing fuel consumption and maintaining the driving comfort. Additionally, it is required that the air-fuel ratio is kept at $R_{\mathrm{a/f}} = 1$ during homogeneous combustion mode. The engine-speed shall always be larger than a specified idle-speed limit to avoid stalling, also in the presence of load disturbances. Minimizing fuel consumption is in principle equivalent to using stratified combustion mode whenever possible. Maintaining driver comfort involves minimizing torque bumps due to mode changes and avoiding too large torque gradients, especially at low gears.

The EMS control signals are the command signal to the fresh-air throttle (MTC), the ignition-advance angle (IGA), the start of injection (SOI), and the fuel-injection duration (T_{inj}), i.e., the amount of fuel injected. The start of injection essentially determines the combustion mode that the engine operates in. The problem of finding optimal IGA and SOI is not part of the benchmark, and suitable values are assumed to be known for all operating conditions. The manifold pressure (P_{man}), throttle position (MTC_{pos}), engine speed (N), and air-fuel ratio ($R_{\text{a/f}}$) are measured signals available for feedback. The air inlet temperature and the atmospheric pressure are measured, but regarded as constant in this study. It is also assumed that the effective torque (TQE) is available as a feedback signal. The feasibility of this assumption is discussed in Section 2 below.

The first stage of the benchmark is to design a core controller that handles torque-tracking and mode-switching, see Figure 3. During this

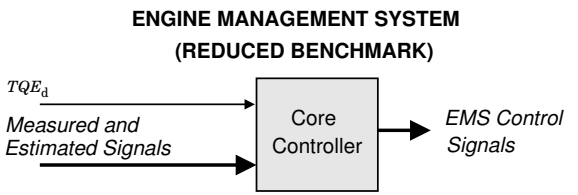

Figure 3 Structure of the Engine Management System in the reduced benchmark.

stage a reduced benchmark model is used, where the driver is excluded. The reference inputs to the EMS are instead effective torque trajectories generated by calculating the corresponding demanded torque from three small excerpts, or scenarios, of the full European driving cycle. The scenarios are chosen such that the gear is constant.

In the second stage of the benchmark study, the goal is to design an engine management system for the full benchmark shown in Figure 2. Instead of a computed torque reference, the command signals to the controllers are now the accelerator, brake, and clutch pedal positions, and the gear. These signals come from a model that mimics the behavior of a human driver. The inputs to the driver model are the reference signals (vehicle speed and gear) from the European driving cycle. The control problem is extended to also include idle-speed control with the presence of electric load disturbances. The complete engine management system is shown in Figure 4.

Torque estimation and sensing

Availability of the effective torque signal TQE relies on either estimation or direct measurements. Real-time estimation of the engine torque

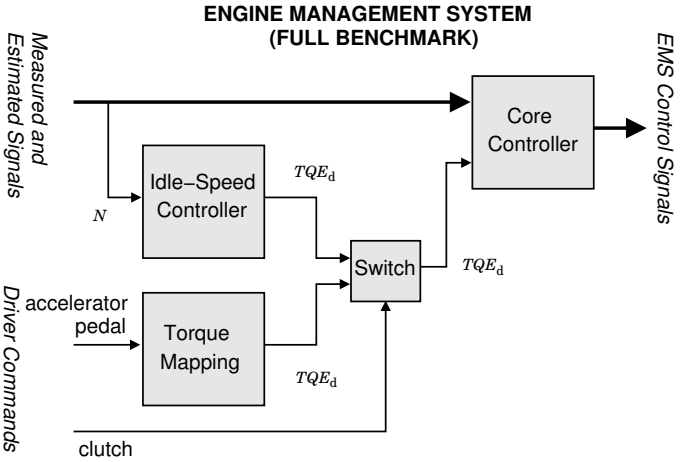

Figure 4 Structure of the Engine Management System in the full benchmark. The core controller is extended with the idle-speed controller, the accelerator pedal interpreter, and switching logic that determines which of these blocks that provides the torque reference.

is not trivial. Earlier work has mainly focused on estimation of indicated torque for diagnosis purposes. With the introduction of Electronic Throttle Control the interest has shifted to construct real-time torque estimators that can act as virtual sensors in closed-loop engine control. In [Rizzoni, 1989] effective torque estimates are obtained from a second-order passive circuit model using crankshaft speed fluctuations. [Azzoni *et al.*, 1998] presents a summary of results concerning the estimation of indicated, effective, and load torques in IC engines using nonlinear second-order sliding mode observers using driveline and crankshaft speed measurements. Other related work is presented in [Moskwa and Pan, 1995, Wang *et al.*, 1997, Gyan *et al.*, 2000, Haskara and Mianzo, 2001]. Torque estimation is in general dependent on sophisticated and accurate engine models. Hence, the possibility to use the torque signal to simplify the design of the engine controller has to be weighted against the additional complexity of the estimator.

Another alternative is to measure the effective torque directly. Magnetoelastic, or magnetostrictive, sensors are a possible solution to this problem as they provide low-cost contact-free measurements suitable for harsh environments. Magnetoelasticity is a property of ferromagnetic materials which causes the magnetic permeability of the material to change when it is subjected to mechanical stress. The basic principle of a magnetoelastic

sensor is to let a magnetic field penetrate the shaft, and detect changes in flux caused by the torsional stress. In [Sobel *et al.*, 1996, Jeremiasson and Wallin, 1998, Wallin *et al.*, 2002] practical application of the ABB Torductor sensor is demonstrated. More work on magnetostrictive torque measurements is found in e.g. [Fleming, 1989]. Even if these sensors are generally not found on current production vehicles, they constitute a realistic and mature solution to torque measurements in automotive vehicles.

3. The engine model

The engine model is derived and implemented in Simulink by Siemens Automotive. The model is a continuous-variable mean-value engine model (MVEM) for one cylinder. MVEMs are generally accepted as the modeling paradigm for engine control, and are extensively described in the literature [Hendricks, 1990, Hendricks, 1997, Sun *et al.*, 1999, Andersson *et al.*, 1999, Hendricks *et al.*, 2000]. Therefore this section will be kept brief, and the reader is referred to the cited references for further details.

The model is built using equations derived from first principles combined with engine maps derived through test-bench experiments on a real GDI engine, see [Passaquay, 2000] and [Passaquay *et al.*, 1998]. Hence, some of the analytical physical relations found in standard MVEMs are parameterized with look-up tables. The basic version of the model does not include EGR. However, in Section 4 it is extended with a simple EGR model. The powertrain model can be divided into two parts: the engine model and the engine-load model, see Figure 5. The engine part can be decomposed into three subsystems: air intake, intake manifold, and cylinder. The engine-load part includes clutch, gear box, differential, brakes, and the chassis. The dynamics of the model can be described using three state variables, and the model has a large number of nonlinearities.

The engine hosts two of the states: the throttle position (MTC_{pos}) and the intake-manifold pressure (P_{man}). The engine is equipped with a controlled electronic air-throttle with dynamics that may be approximated by a first order LTI system. The throttle position MTC_{pos} is thus described by

$$T_{\mathrm{MTC}}(d/dt)MTC_{\mathrm{pos}} + MTC_{\mathrm{pos}} = MTC \tag{1}$$

The flow of intake air is controlled with the desired throttle-position signal MTC. The air mass-flow rate through the throttle into the intake manifold (Q_{TPS}) is described by a calibrated look-up table parameterization of the standard orifice equation as

$$Q_{\mathrm{TPS}} = \Psi_0\left(MTC_{\mathrm{pos}}, P_{\mathrm{man}}/P_{\mathrm{atm}}\right) \tag{2}$$

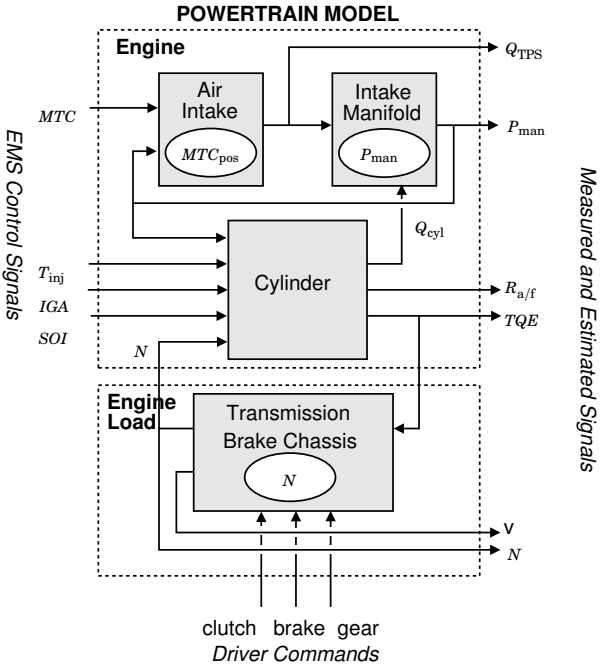

Figure 5 The powertrain model is composed of the engine model and the engine load model. The engine model hosts the throttle position state MTC_{pos} and the intake-manifold pressure state P_{man}. The engine load model hosts the engine-speed state N.

The build-up of the manifold pressure P_{man} is described by the ideal gas law together with a mass-flow balance equation

$$(d/dt)P_{\text{man}} = K_P\big(Q_{\text{TPS}} - Q_{\text{cyl}}\big) \tag{3}$$

The air-flow into the cylinder (Q_{cyl}) depends on the engine speed, manifold pressure and the volumetric efficiency, and is described by a look-up table

$$Q_{\text{cyl}} = \Psi_7\big(P_{\text{man}}, N\big) \tag{4}$$

The torque production in the cylinder is described by a nonlinear mapping that depends on air-flows, engine speed, and the engine control variables T_{inj}, SOI and IGA. The effective torque is the indicated torque generated by the cylinder (TQI), subtracted by the torque losses due to pumping work and friction (TQL).

$$TQE = TQI - TQL \tag{5}$$

The torque losses depend on the engine speed and the air-flow and are described by a look-up table:

$$TQL = \Psi_8(Q_{\text{cyl}}, N) \tag{6}$$

In stratified mode torque production is modeled by the following nonlinear maps.

$$\begin{aligned} TQI &= TQI_{\text{S,ref}} G_{\text{S,phase}} G_{\text{S,air}} \\ TQI_{\text{S,ref}} &= \Psi_{10}(M_{\text{fuel}}, N) \\ G_{\text{S,phase}} &= \Psi_{11}\Big(IGA - \Psi_{12}(TQI_{\text{S,ref}}, N), EOI - IGA - \Psi_{13}(TQI_{\text{S,ref}}, N)\Big) \\ G_{\text{S,air}} &= \Psi_{34}(\Psi_{33}(TQI_{\text{S,ref}}, N) - M_{\text{air}}) \end{aligned} \tag{7}$$

The indicated torque TQI is modeled as the ideal torque $TQI_{\text{S,ref}}$ for a certain fuel-flow and engine speed, penalized by $G_{\text{S,phase}}$ for inappropriate ignition timing, and by $G_{\text{S,air}}$ for inappropriate air-flow. In homogeneous mode the corresponding maps are

$$\begin{aligned} TQI &= TQI_{\text{H,ref}} G_{\text{H,phase}} G_{\text{H,air/fuel}} \\ TQI_{\text{H,ref}} &= \Psi_{32}(M_{\text{air}}, N) \\ G_{\text{H,phase}} &= \Psi_{15}\Big(IGA - \Psi_{16}(M_{\text{air}}, N)\Big) \\ G_{\text{H,air/fuel}} &= \Psi_{17}(R_{\text{a/f}} - \Psi_{18}(M_{\text{air}}, N)) \end{aligned} \tag{8}$$

The indicated torque TQI is modeled as the ideal torque $TQI_{\text{H,ref}}$ for a certain air-flow and engine speed, penalized by $G_{\text{H,phase}}$ for inappropriate ignition timing, and by $G_{\text{H,air/fuel}}$ for inappropriate air-fuel ratio.

Together with the engine speed N the manifold pressure determines the amount of air taken into the cylinder, Q_{air}, while T_{inj} and SOI determine the amount of fuel injected and the injection timing respectively. Note that the cylinder subsystem contains no dynamics. The outputs TQE and $R_{\text{a/f}}$ are thus momentarily determined by the inputs IGA, SOI, T_{inj}, P_{man}, and N. The third state, the engine speed N, is hosted by the transmission-brake-chassis subsystem, and varies slowly compared to the dynamics of the engine subsystem. It is described by

$$J(d/dt)N = TQE - TQR \tag{9}$$

where TQR depends on the gear, and includes rolling-resistance, aerodynamic drag, climbing-resistance, and brake resistance. Since it is not an objective to control N, it will be regarded as an external variable on the engine subsystem in the core-controller design.

4. Controller design

The engine is a highly nonlinear multi-input multi-output system. To handle the complexity a bottom-up approach has been adopted. A number of specialized simple controllers for the different subsystems of the motor have been designed, and then integrated into a complete control system.

Considerations on model uncertainties and tuning issues led to a controller structure based on extensive use of feedback-loops with simple feedforward paths. In combination with the bottom-up approach called for by the complexity, this led to a structure based on several sub-controllers. The sub-controllers can be tuned individually, with only rough process-knowledge, and with only a few parameters to calibrate.

Choice of control variables

In the context of torque control the cylinder operation can be described as nonlinear static maps, constructed by aggregating the look-up tables Ψ_i of the previous section

$$\begin{aligned}(T_{\text{inj}}, P_{\text{man}}, N, IGA, SOI) &\overset{\text{Homogeneous}}{\mapsto} TQE \\ (T_{\text{inj}}, P_{\text{man}}, N, IGA, SOI) &\overset{\text{Stratified}}{\mapsto} TQE\end{aligned} \tag{10}$$

The variables IGA and SOI are chosen to yield optimal efficiency. Since N varies slowly it is then sensible to consider a family of maps

$$\begin{aligned}(P_{\text{man}}, T_{\text{inj}}) &\overset{\text{Homogeneous}}{\underset{N}{\mapsto}} TQE \\ (P_{\text{man}}, T_{\text{inj}}) &\overset{\text{Stratified}}{\underset{N}{\mapsto}} TQE\end{aligned} \tag{11}$$

that are parameterized in N. These maps describe how the generated effective torque TQE depends on the manifold pressure P_{man} and the fuel injection time T_{inj} at given engine speeds N. Experiments on the Simulink model of the cylinder were used to derive these maps. Figure 6 shows the maps for $N = 1500$ rpm.

Understanding the characteristics of these maps is an essential prerequisite for successful torque control. For example, from the maps it is easy to see that the sensitivity of TQE to variations in P_{man} is large in homogeneous mode while it is large to variations in T_{inj} in stratified mode.

Recall from the previous section that T_{inj} can be set instantaneously, while P_{man} is the result of a dynamic response to MTC. In homogeneous mode the $R_{\text{a/f}}$ constitutes the most critical controlled variable. Therefore T_{inj} is chosen as control variable for the $R_{\text{a/f}}$ in this mode. The variable

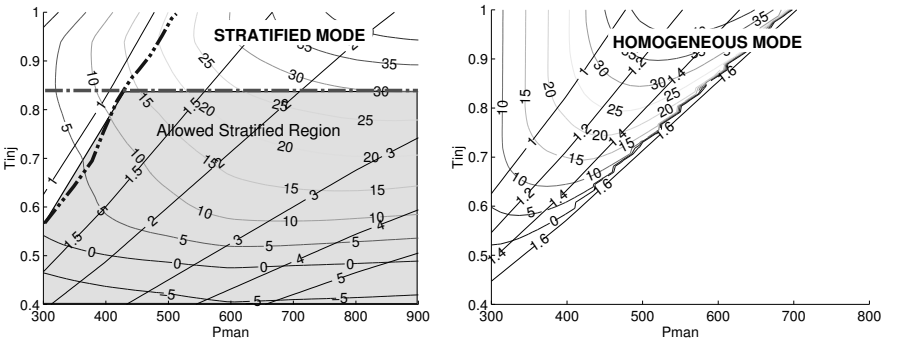

Figure 6 Maps of the cylinder characteristics at $N = 1500$ rpm in stratified mode (left) and homogeneous mode (right). The plots show constant-torque level-curves as a function of the manifold pressure P_{man} and the injection duration T_{inj}, with *IGA* and *SOI* chosen optimally. The diagonal straight lines show constant $R_{\text{a/f}}$. The shaded area shows the allowed region for stratified mode. The dash-dotted line marks the constraint on $TQI_{\text{S,ref}}$, and the dash-dot-dotted line the constraint on M_{air}.

P_{man} is then used to control TQE. In stratified mode the only requirement on the air-fuel ratio is $R_{\text{a/f}} \geq 1$. Thus T_{inj} is the natural choice for controlling TQE, and P_{man} can be chosen such as to yield the $R_{\text{a/f}}$ that minimizes the fuel consumption. (It is also evident from Figure 6 that the torque sensitivity to changes in P_{man} is small for lean air-fuel ratios). It can be seen in Figure 6 that there exists an optimal combination of P_{man} and T_{inj} for obtaining a specific torque TQE from the engine in stratified mode. This choice minimizes T_{inj}, which is equivalent to minimizing instant fuel-consumption.

Manifold pressure controller

In [Kume *et al.*, 1996] it is stated that a quick and accurate control of the air-flow is necessary to ensure smooth transitions between stratified and homogeneous mode. Since the air-flow is closely related to the manifold pressure it was therefore decided to control the air-flow indirectly by controlling P_{man}.

The manifold-pressure dynamics (3) can be approximated by the differential equation

$$(d/dt)P_{\text{man}} \approx k_1 MTC_{\text{pos}} f(\eta) - k_2 N P_{\text{man}} \tag{12}$$

where $\eta = P_{\text{man}}/P_{\text{atm}}$. From the look-up table relating P_{man} and Q_{TPS} it was heuristically found that this nonlinearity can be approximated with $\hat{f}(\eta) = 1 - \eta^8$.

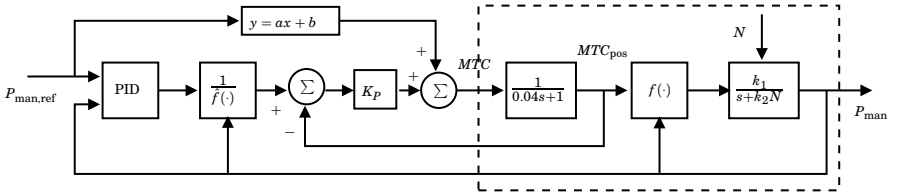

Figure 7 Block diagram showing throttle and manifold-pressure controllers and the corresponding dynamics. Blocks inside the dashed box are in continuous time.

A proportional control of MTC_{pos} is first applied to obtain a desired time constant of the fresh-air throttle.

$$MTC(k) = K_P\,[MTC_{\text{pos,d}}(k) - MTC_{\text{pos}}(k)] \tag{13}$$

where $MTC_{\text{pos,d}}$ is a desired throttle-position. A PID controller is then introduced to control P_{man} to follow the desired trajectory $P_{\text{man,ref}}$. To eliminate the gain variations for the P_{man} dynamics, the control signal is scaled as

$$MTC = \frac{u}{\hat{f}(\eta)} \tag{14}$$

where u is the output from the PID-controller for controlling P_{man}. The $\hat{f}(\eta)$ term can be viewed as a local feedback linearization term. Assuming that $MTC_{\text{pos}} \approx MTC$, the resulting differential equation, seen from the PID-controller, is

$$(d/dt)P_{\text{man}}(t) \approx k_1 u(t) - k_2 N(t) P_{\text{man}}(t) \tag{15}$$

Note that the open-loop time-constant from P_{man} still varies with N. Since N is measured a possibility would be to use gain scheduling. Instead the PID-controller was robustly tuned using the Nonlinear Control Design Blockset in Simulink, to handle the variations in N. The resulting closed-loop system has a time constant of around 60 ms, which appears constant in the entire work space. The resulting P_{man} controller, hence, is insensitive to variations in N. A feedforward term corresponding to an affine approximation of the static relation from P_{man} to MTC_{pos} was added to speed up the response to changes in $P_{\text{man,ref}}$. The result is the cascade feedback controller-structure with feedback linearization and feedforward in Figure 7.

Torque controller in stratified mode

As stated above, T_{inj} is used for controlling TQE in stratified mode. Since the references for TQE contain ramp variations it is necessary to have

two integrators in the controller to achieve good tracking. Therefore a PII^2-controller was used:

$$C(q) = K_p + \frac{K_i h}{q-1} + \frac{K_{i^2} h^2}{(q-1)^2} \tag{16}$$

In stratified mode a rough linear approximation of the torque generation map is $\delta TQE = (60 \pm 30)\delta T_{\text{inj}} + (0.02 \pm 0.02)\delta P_{\text{man}}$. With the unit delay introduced by the computational delay in the controller a rough model is

$$\delta TQE \approx \frac{60 \pm 30}{q} \delta T_{\text{inj}} \tag{17}$$

The resulting closed-loop system is of third order. A pole-placement design assuming a nominal gain of 60 gives satisfactory performance and robustness.

A feedforward term corresponding to an affine approximation of the static relation from T_{inj} to TQE is added to speed up the response to reference changes.

Torque controller in homogeneous mode

In homogeneous mode a rough linear approximation of the torque generation map is $\delta TQE \approx (0.1 \pm 0.05)\delta P_{\text{man}}$ (with P_{man} and T_{inj} constrained by the requirement that $R_{\text{a/f}} = 1$). Combining this with a linear first-order system approximation of the controlled manifold pressure dynamics gives

$$\delta TQE \approx \frac{(0.10 \pm 0.05)(1 - e^{-h/0.06})}{q - e^{-h/0.06}} \delta P_{\text{man,ref}} \tag{18}$$

where h is the sampling time. A PII^2-controller was tuned using pole placement, assuming a nominal gain of 0.1. It was necessary to also introduce a lead compensator to obtain a phase gain around the cross-over frequency. The torque control is considerably slower in homogeneous mode than in stratified mode, due to the controlled manifold-pressure dynamics present in the loop.

A feedforward term corresponding to an affine approximation of the static relation from P_{man} to TQE was added to speed up the response to reference changes.

Air-fuel ratio controller in homogeneous mode

The air-fuel ratio control problem in homogeneous mode is not the main focus of this work. Therefore it is assumed that a well working controller exists, approximated by controlling $R_{\text{a/f}}$ using T_{inj}, with a pure feedforward strategy based on the stoichiometric ratio. This control yields perfect control of $R_{\text{a/f}}$ with a delay of one sample.

Extremum seeking

Extremum control of $R_{a/f}$**.** In stratified mode the torque TQE is regulated with the control variable T_{inj}. The $R_{a/f}$ can be chosen arbitrarily. In [Iwamoto *et al.*, 1997] it is suggested to use a constant ultra-lean air-fuel ratio in stratified mode, while in [Jackson *et al.*, 1997] light throttling is suggested to reduce emissions. In the present work the $R_{a/f}$ is instead chosen as to maximize engine efficiency, which is equivalent to choose P_{man} to minimize fuel consumption. Thus for constant TQE and N the engine efficiency can be described by the fuel consumption $T_{inj} = g(P_{man})$, see Figure 6. The minimum of this function is dependent on operating conditions such as N and TQE, but also on e.g. aging effects on the engine. An extremum controller can be used to find and track this minimum during engine operation, without any other prior knowledge on the shape of g than an assumption of convexity. The controller applies a perturbation δP_{man} to the base manifold pressure $P_{man,base}$. The effect of the perturbation on the engine output-torque TQE is eliminated by closed-loop control using T_{inj}. In this way the effects of the perturbations are only visible internally in the controller as correcting variations ΔT_{inj} in the fuel mass-flow. The suggested structure with the interconnection of an extremum controller and a controller that eliminates the effects of the perturbation signals on the process output is therefore denoted a *silent* extremum controller. To the authors knowledge this is a new construction. The gradient direction of the engine efficiency with respect to the manifold pressure is computed from the perturbation and the corresponding variation in the fuel mass-flow. The base manifold-pressure is then changed in the negative gradient direction. The control structure is shown in Figure 8, and may be implemented as

$$P_{man,base}(k) = P_{man,base}(k-1) - \gamma(|\Delta T_{inj}(k)|)\,\mathrm{sgn}\left(\frac{\Delta T_{inj}(k)}{\Delta P_{man}(k)}\right) + \Delta\delta P_{man}(k)$$

where

$$\begin{aligned}\Delta P_{man}(k) &= P_{man}(k) - P_{man}(k-1)\\ \Delta T_{inj}(k) &= T_{inj}(k) - T_{inj}(k-1)\\ \Delta\delta P_{man}(k) &= \delta P_{man}(k) - \delta P_{man}(k-1)\end{aligned}$$

where $\gamma(\cdot)$ is a function chosen to give an appropriate step length. The perturbation signal is chosen as $\delta P_{man} = A \sin \omega t$. The consecutive corrections make the base manifold pressure converge to a value corresponding to the optimal $R_{a/f}$ for the present operating conditions.

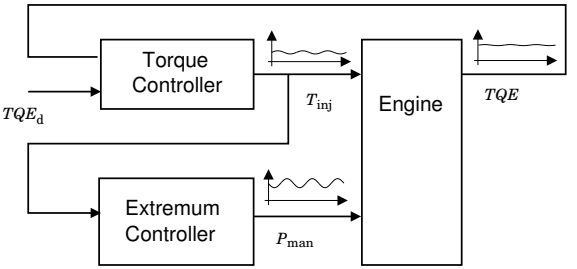

Figure 8 Structure of the silent extremum control strategy.

Note that the extremum-controller does not use any information about the engine, but only relies on measurements collected during normal engine operation. The silence property of the controller is critically dependent on the closed-loop torque control. It would be possible to use control of other engine outputs such as engine speed instead. A review of extremum control can be found in [Sternby, 1979]. Some approaches to off-line optimization of GDI engines are presented in [Kolmanovsky *et al.*, 1999].

Extremum control of $R_{a/f}$ and EGR. In [Sasaki *et al.*, 1998] it is reported how EGR reduces fuel consumption and NO_x emissions in a GDI engine running in stratified mode. An interesting feature of the results is that the fuel consumption attains its minimum at the desired EGR operation point regarding pollutant emissions. Therefore extremum control of $EGR = Q_{EGR}/(Q_{TPS} + Q_{EGR})$ with minimization of T_{inj} as control objective may be feasible. The recirculated exhaust-gas mass-flow Q_{EGR} is controlled with the EGR-valve command signal ETC. If it is desired to optimize fuel consumption with respect to both P_{man} and EGR simultaneously it is necessary to modify the previous approach, such that the extremum controller can separate the effects of the perturbation signals on P_{man} and EGR on the engine output, i.e., in effect on T_{inj}. This may be achieved by choosing orthogonal perturbation signals. The effect of small deviations δP_{man} and δETC from a base manifold-pressure $P_{man,base}$ and a base EGR-valve position ETC_{base} can then be described by the first order approximation

$$T_{inj} \approx T_{inj}|_{base} + \left.\frac{\partial T_{inj}}{\partial P_{man}}\right|_{base} \delta P_{man} + \left.\frac{\partial T_{inj}}{\partial ETC}\right|_{base} \delta ETC$$

With $\delta P_{man} = a_P \sin(m_P \omega t)$ and $\delta ETC = a_E \sin(m_E \omega t)$ with integers $m_P \neq m_E$, it follows after multiplication of the expression above with a

perturbation signal, and integration over k periods, that the gradient can be computed as

$$\left.\frac{\partial T_{\text{inj}}}{\partial P_{\text{man}}}\right|_{\text{base}} = \frac{m_P \omega}{k a_P^2 \pi} \int_{t-k2\pi/\omega}^{t} \delta P_{\text{man}}(s) T_{\text{inj}}(s)\, ds$$

$$\left.\frac{\partial T_{\text{inj}}}{\partial ETC}\right|_{\text{base}} = \frac{m_E \omega}{k a_E^2 \pi} \int_{t-k2\pi/\omega}^{t} \delta ETC(s) T_{\text{inj}}(s)\, ds$$

After an integration period, $P_{\text{man,base}}$ and ETC_{base} are updated with a step in the negative gradient direction. This corresponds to a steepest-descent optimization, which is known to have slow convergence. A possible refinement would be to estimate also the Hessian matrix, and use more elaborate methods to compute the step.

To evaluate the extremum controller on combined EGR and P_{man} optimization the calibrated engine model was extended with a simple isothermal EGR model [Hendricks *et al.*, 2000], where the EGR flow is described by the standard orifice equation. A more accurate EGR model for transient behavior would be obtained by using an adiabatic manifold pressure model as in [Fons *et al.*, 1999], but for the purposes of this study the isothermal model is considered to be sufficient. An efficiency map derived from the results in [Sasaki *et al.*, 1998] was added to the torque production model. For a certain TQE, N and EGR an efficiency factor on the indicated torque was computed by multi-linear interpolation of the experimental data. This is believed to yield at least a crude approximation of real engine behavior. Moreover, the results of the extremum controller are valid even though the real efficiency dependence would deviate from the model, since the lack of a priori knowledge of the exact shape of the target function is one of the distinguishing features of extremum control. The resulting efficiency maps are shown in Figure 9.

Idle-speed control

The objective of the idle-speed controller is to keep the engine speed at a nominal speed N_{idle}. Undershoots below the stalling limit N_{stall} are not permitted. Load disturbances may be present due to electrical equipment being turned on or off. Large load disturbances are preceded by a warning signal to the controller, indicating that a large load is pending. The controller may delay the arrival of the large load for a short while. This can be used to prepare the idle-speed controller for the forthcoming load, and is a common solution in modern vehicles.

In line with the simple controller structures used in the core controller the idle-speed controller has also been kept simple. It consists of a simple

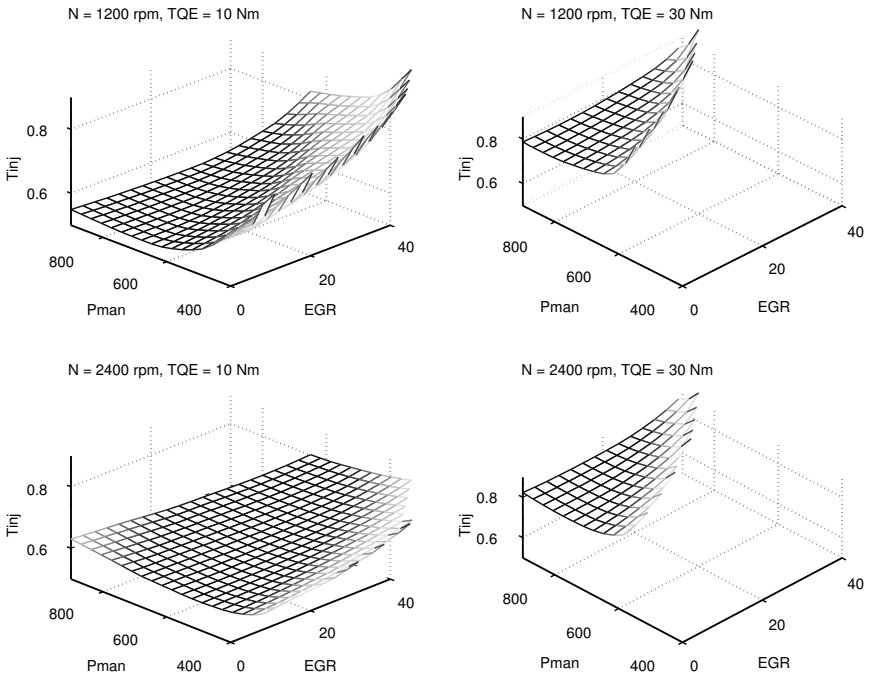

Figure 9 Maps of cylinder characteristics regarding fuel consumption T_{inj}, showing how P_{man} and EGR can be chosen optimally to yield a certain desired TQE for a given N.

PI-controller where the proportional gain depends slightly on the engine speed. The input to the idle-speed controller is the idle-speed error, i.e, $N_{\text{idle}} - N$, and the output is the torque reference. The warning signal for large loads is used to add a constant value to the integrator state of the PI-controller. This is a simple solution that prevents any dangerous undershoots in the engine speed when the large load arrives.

Switching strategies

An elaborate switching-strategy is used for switching between homogeneous and stratified mode. The mode-switches should ideally be performed without large torque-bumps. The strategy tries to switch mode where the torque level curves in homogeneous mode and stratified mode meet. To achieve this the $R_{\text{a/f}} = 1$ constraint in homogeneous mode is temporarily relaxed during the mode-switch.

The engine should run in stratified combustion mode whenever pos-

sible. The controller thus has to detect when conditions for stratified operation are fulfilled, and then perform a mode-switch. When the constraints are no longer fulfilled the combustion mode immediately has to be switched back to homogeneous. The constraints for running in stratified mode have in the benchmarks been defined as:

- $TQI_{\mathrm{S,ref}} < 50$ (essentially a bound on T_{inj})
- M_{air} is large enough (essentially a bound on P_{man})
- $N < 3000$

The constraint limits of the physical engine are confidential, therefore the model constraints are set to typical values.

Only the first constraint poses a real challenge. Since T_{inj} varies quickly at fast torque-reference changes, the constraint may be violated with short notice. To handle such situations a simple $TQI_{\mathrm{S,ref}}$ predictor was implemented as

$$\widehat{TQI}_{\mathrm{S,ref}}(kh + t) = TQI_{\mathrm{S,ref}}(kh) + (TQE_{\mathrm{d}}(kh) - TQE_{\mathrm{d}}(kh - h))T/h \quad (19)$$

With a prediction horizon $T = 0.2$ s the need for mode switches can be detected in advance, and constraint violations can be avoided.

The mode-switches should ideally be performed without large torque bumps. By superimposing the torque level curves of Figure 6 the plot in Figure 10 is obtained. In this plot it is observed that there exist points in the $(P_{\mathrm{man}}, T_{\mathrm{inj}})$-space where the torque difference between homogeneous and stratified mode is zero. A key observation is that these points lie close to a line described by a constant air-fuel ratio close to the lean limit, $R_{\mathrm{a/f}} \approx 1.5$. The mode-switch strategy adopted in this work is to try to perform all combustion-mode switches on this line. Descriptions of similar mode-switch strategies are described in [Iwamoto *et al.*, 1997, Kume *et al.*, 1996, Zao *et al.*, 1997, Druzhinina *et al.*, 1999].

The mode switching strategy is best illustrated by paths in a torque-level diagram as in Figure 11. In this diagram lines of constant torque in homogeneous and stratified mode are plotted, together with lines of constant air-fuel ratio, stratified-mode operation constraints, and the line representing zero torque difference.

In the following subsections the switching strategies between the stratified and homogeneous modes are described in some detail. The mode switching logic is modeled as a state machine that is implemented using Simulink Stateflow™. The letters used in the descriptions below refer to the paths in the figure.

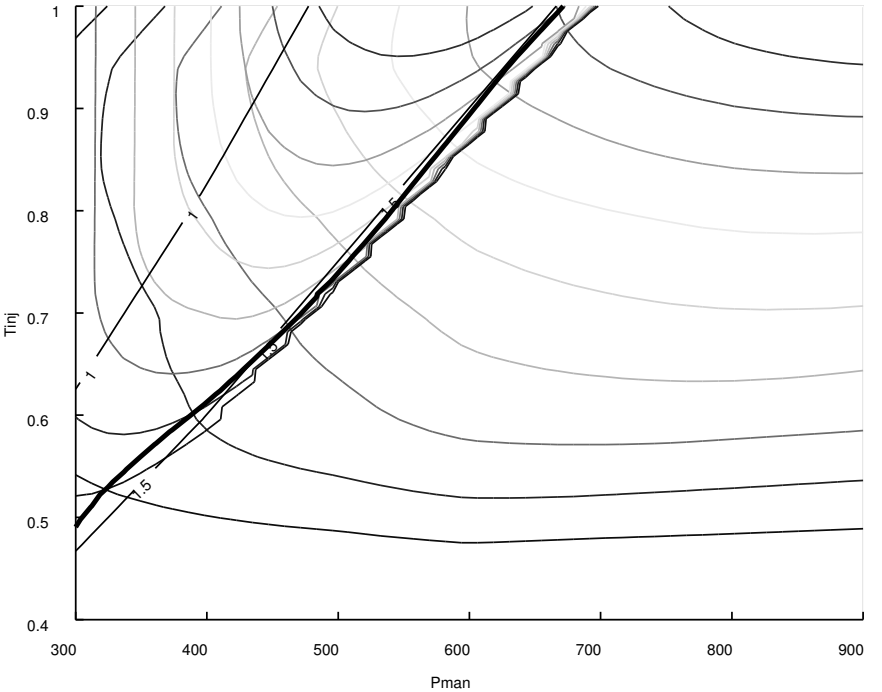

Figure 10 Static cylinder map $(P_{\text{man}}, T_{\text{inj}}) \mapsto TQE$ for stratified and homogeneous mode. The bold line indicates a set of points where the torques in homogeneous and stratified mode are equal.

Switch from stratified to homogeneous mode. This mode-change poses the greatest challenge because of the constraint on $TQI_{\text{S,ref}}$ to be upper bounded in stratified mode. A mode-switch is initiated when $TQI_{\text{S,ref}}$ is predicted to violate the stratified constraints within 0.2 s. The mode-switch consists of a number of submodes:

$A \longrightarrow B$**:** In this step P_{man} is decreased as quickly as possible to reach the zero torque-difference line before the $TQI_{\text{S,ref}}$ limit is reached (torque is probably increasing rapidly). The extremum controller is switched off, and a reference value $P_{\text{man,ref}}$ corresponding to $R_{\text{a/f}} = 1.5$ is passed to the P_{man}-controller.

$B \longrightarrow C$**:** If the $TQI_{\text{S,ref}}$ limit is reached T_{inj} has to be kept constant until $R_{\text{a/f}}$ has reached the zero torque-difference line. At this point the TQE-controller is switched off, and T_{inj} is frozen to its present value. Therefore as P_{man}, and in effect $R_{\text{a/f}}$, decreases, the torque error

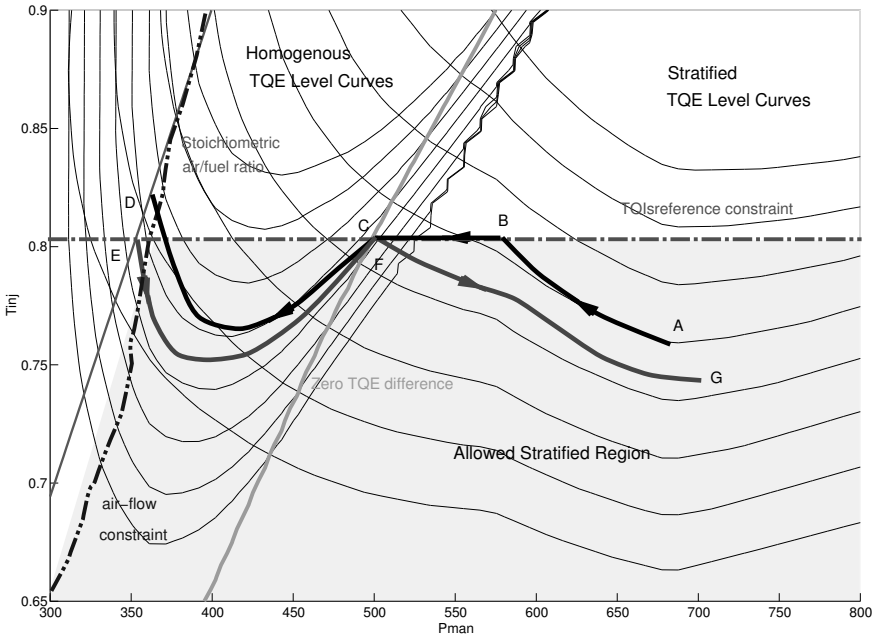

Figure 11 Illustration of the switch strategy between homogeneous and stratified mode. The path $A \rightarrow B \rightarrow C \rightarrow D$ illustrates a switch from stratified to homogeneous mode where the torque limit of the stratified mode is reached before the switch is ended. The mode switch occurs at C and the switch of control strategy occurs at D. The path $E \rightarrow F \rightarrow G$ illustrates a switch from homogeneous to stratified mode. Here the control-strategy switch occurs at E and the mode-switch occurs at F.

increases.

C**:** The actual mode-switch takes place close to the zero torque-difference line $R_{\mathrm{a/f}} \approx 1.5$.

$C \longrightarrow D$**:** After the mode switch P_{man} is used to regulate $R_{\mathrm{a/f}}$ to its nominal reference in homogeneous mode. The torque is controlled with T_{inj}.

D**:** When $R_{\mathrm{a/f}}$ is close to the nominal reference the homogeneous-mode controllers are switched on, which use T_{inj} to control the air-fuel ratio, and P_{man} to control the torque.

Switch from homogeneous to stratified mode. A mode-switch from homogeneous to stratified mode is triggered when the constraints for run-

ning in stratified mode are fulfilled. This switch does not impose as critical timing conditions as the switch in the opposite direction, since the torque is decreasing, and $TQI_{\mathrm{S,ref}}$ is moving away from its bound in stratified mode. A certain margin is required regarding $TQI_{\mathrm{S,ref}}$ to avoid constraint violations during the mode-switch, and to avoid a situation with instant chattering mode switches.

$E \longrightarrow F$**:** A controller that regulates the torque with T_{inj} is switched on, while P_{man} is ramped up to the zero torque-difference line.

$F \longrightarrow G$**:** The actual mode switch takes place close to the zero torque-difference line. P_{man} is ramped up to a suitable starting position for the extremum controller.

G**:** The extremum-seeking controller is switched on.

The mode-switching strategy described above temporarily, during the switching, relaxes the constraint that the air-fuel ratio should be equal to 1 in homogeneous mode. This increases the amount of emissions slightly. Since the benchmark does not include an emissions model it is difficult to evaluate the severity of these additional emissions.

Complete controller structure

Reduced benchmark. The resulting controller structure is shown in Figure 12. The core controller consists of five sub-controllers. The P_{man} controller is central in the total controller design. In stratified mode the extremum controller minimizes the fuel consumption by producing a suitable reference to the P_{man} controller. The *TQE* controller controls the effective torque using T_{inj}. In homogeneous mode, the *TQE* controller controls the effective torque by producing a suitable reference to the P_{man} controller. A simple controller regulates the $R_{\mathrm{a/f}}$ in homogeneous mode using T_{inj}. The control signals *IGA* and *SOI* are obtained from maps. The mode-switches are controlled by a discrete-event logic controller.

The sub-controllers that are used in either homogeneous mode, stratified mode, or during mode switches, have to be switched in and out in a graceful manner. To ensure this an extensive *tracking-scheme* is included in the core controller. The same tracking-scheme is used to prevent integrator-windup when the actuators saturate.

The core controller is a discrete-time controller with sampling period 5 ms. Anti-aliasing filters were added to the measured signals. Please refer to [Åström and Wittenmark, 1997] for more details on design and implementation of dicrete-time controllers.

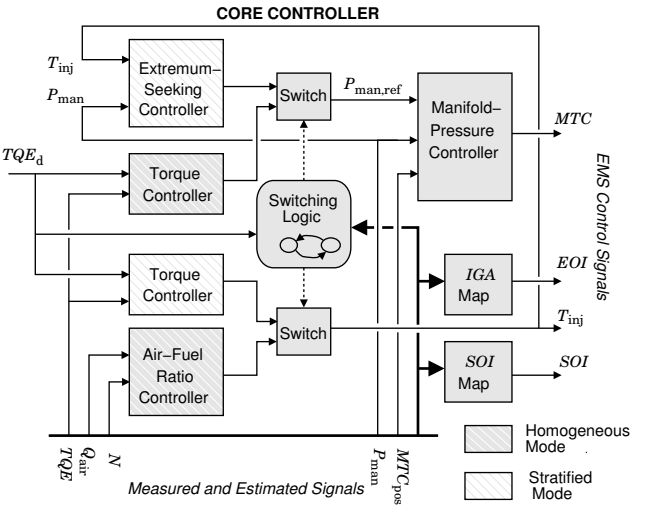

Figure 12 Overall structure of the sub-controllers in the core controller.

Full benchmark. The controller structure for the complete engine management system is shown in Figure 4. The accelerator-pedal position is mapped to a corresponding torque reference using a simple linear transformation where the accelerator-pedal range $0 - 100\%$ is mapped to the range $TQE_{\min}(N) - TQE_{\max}(N)$, where the limits depend on the engine speed. The idle-speed controller is switched in when the accelerator pedal is fully released, or when the clutch is fully disengaged and the torque reference generated by the idle-speed controller is larger than the torque reference generated by the driver. The latter part of the condition is used to prevent engine stalling during the switch from idle-speed control to driver control during gear switches. A rate limiter is used on the torque reference to improve the performance during gear-switches. To avoid unnecessary mode-changes it is not possible to switch from homogeneous mode to stratified mode during a gear-switch (when the clutch is disengaged).

The resulting control structure is of multi-level cascade type. The outermost level is the driver or the idle speed controller. The next level is the torque controller in the core controller. In homogeneous mode the torque controller generates the reference for the P_{man}-controller which is also a cascade controller. Hence, in the extreme case the system contains four controller levels.

5. Results

Reduced benchmark

The core controller was tested in Simulink simulations with the reduced-benchmark model. The tests include three different scenarios of different durations and torque references, and six different engine configurations:

1. The nominal model without measurement noise.
2. The nominal model with a unmodeled 10% torque loss to simulate aging phenomena and engine-to-engine variations. Without measurement noise.
3. The nominal model with a unmodeled 10% torque gain to simulate engine-to-engine variations. Without measurement noise.
4. The nominal model with measurement noise.
5. Configuration 2 with measurement noise.
6. Configuration 3 with measurement noise.

The main demands of the control system are that the torque should be within ± 2 Nm from the reference during normal operation, that torque-bumps at the mode-switches should be less than 10 Nm and that the air-fuel ratio $R_{\mathrm{a/f}}$ should be controlled with an accuracy within 2%.

The presented core controller gave good performance in simulations, well within the specifications. It was compared with other more traditional controllers that used feedforward from detailed inverse engine maps, and competed well in tracking performance. The extensive use of feedback also made the present controller more robust to engine variations. The torque-gain variations did not affect tracking or mode-switch performance significantly. The controller worked well also in presence of noise, even though noisy signals were not explicitly taken into account in the design procedure. Figure 13 shows the noise-free simulations of the torque response in a torque scenario at gear 1.

The mode-switch from homogeneous to stratified mode takes place at $t \approx 2.6$ s. The simulations show that the torque is followed with good accuracy and that the mode-switch is carried out practically without bumps. The price paid for the small torque-bump is that the $R_{\mathrm{a/f}}$ deviates more from the reference than specified during the transition phase in the mode-switch.

The periodic variations that are present in the MTC, P_{man} and T_{inj} signals in stratified mode, are due to the perturbation signal applied by the extremum controller. The perturbation signal was chosen as $\delta P_{\mathrm{man}} =$

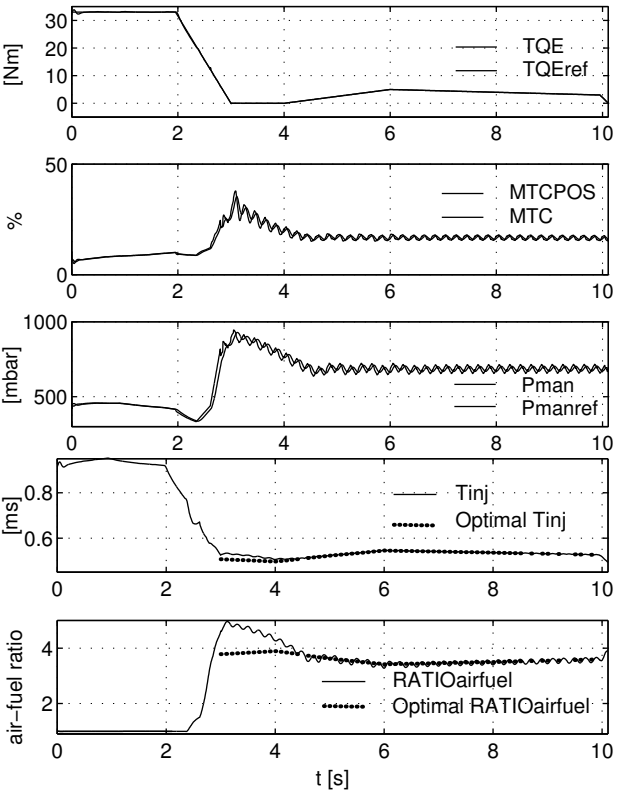

Figure 13 Results for torque scenario at low gear. Stratified mode is entered and the extremum controller is started at $t = 2.6$ s. Note how the perturbation signal is visible as a periodic variation in MTC and P_{man}, but invisible in TQE. The extremum controller converges to minimal T_{inj} in approximately 1.5 s, and is then able to track the optimum. The optimal values are computed by means of numerical optimization on the simulation model.

$30\sin(10\pi t)$ mbar, and $P_{\text{man,base}}$ was updated every 50 ms. Note that the effect of the perturbation signal is not visible in TQE. The controller correctly finds the optimal P_{man} with a convergence time of about 1.5 s. In a real engine the presence of periodic pumping fluctuations in the manifold pressure [Hendricks *et al.*, 1994] that interferes with the perturbation signals may have a deteriorating effect on the extremum controller. A remedy would be to use the controller of Section 4, where the perturbation signal may be chosen orthogonal to the pumping fluctuations.

Results with EGR. The extremum control of *EGR* was evaluated on simulations on the full benchmark extended with the EGR model as described above. In Figure 14 is shown how the extremum controller performs when optimizing P_{man} and *EGR* simultaneously. The engine enters stratified mode at $t = 10.5$ s. The P_{man} extremum controller is started directly when entering stratified mode, while the *EGR* extremum controller is started at $t = 17$ s. Convergence is somewhat slower than in the previous case when only P_{man} is optimized. This can partly be explained by the inefficient steepest-gradient step in the optimization. The perturbation signals was chosen as $\delta ETC = 0.5\sin(2\pi t)$ and $\delta P_{\text{man}} = 10\sin(3\pi t)$. Integration is done over one period, and ETC_{base} and $P_{\text{man,base}}$ are updated every 1 s.

Notice how the effect of the perturbation signals is completely invisible in *TQE*. The perturbation signals were chosen more or less arbitrarily, and the indicated performance of the controllers may be improved further with more careful choices.

Full benchmark

In the full benchmark the evaluation scenario consists of the full European driving cycle including gear switches and idle-speed phases. The core controller without EGR control is used together with the idle-speed controller described in the previous section. The reason for not incorporating EGR in these evaluations is that EGR was introduced to the benchmark at a late stage. Therefore the EGR was not present in the evaluation criteria, and was not part of the solutions of all participating research groups. In the reduced benchmark it was important to achieve good torque-tracking performance. This resulted in the additional I^2 terms in some of the PI-controllers. In the full benchmark, torque tracking is no longer of major importance, since the torque reference is now an internal signal in the EMS. As a result of this it is probably possible to reduce the complexity of the core controller even further.

The engine speed and the vehicle speed for the full driving cycle are shown in Figure 15. The idle-speed has been defined as $N_{\text{idle}} = 1000$ rpm and the engine speed at which the engine starts misfiring has been chosen as $N_{\text{stall}} = 900$ rpm. To evaluate the idle-speed controller, load disturbances l corresponding to electrical loads are introduced during the idle-speed control parts of the driving cycle. The scenario contains two types of loads. Small loads are modeled as non-measurable constant loads of size 3.5 Nm. Large loads, corresponding, e.g., to the air-conditioning system, are also modeled as non-measurable constant loads, now of size 15 Nm.

A typical gear-shift phase is shown in Figure 16. The accelerator pedal is released during the gear-switch. A combustion-mode switch from homo-

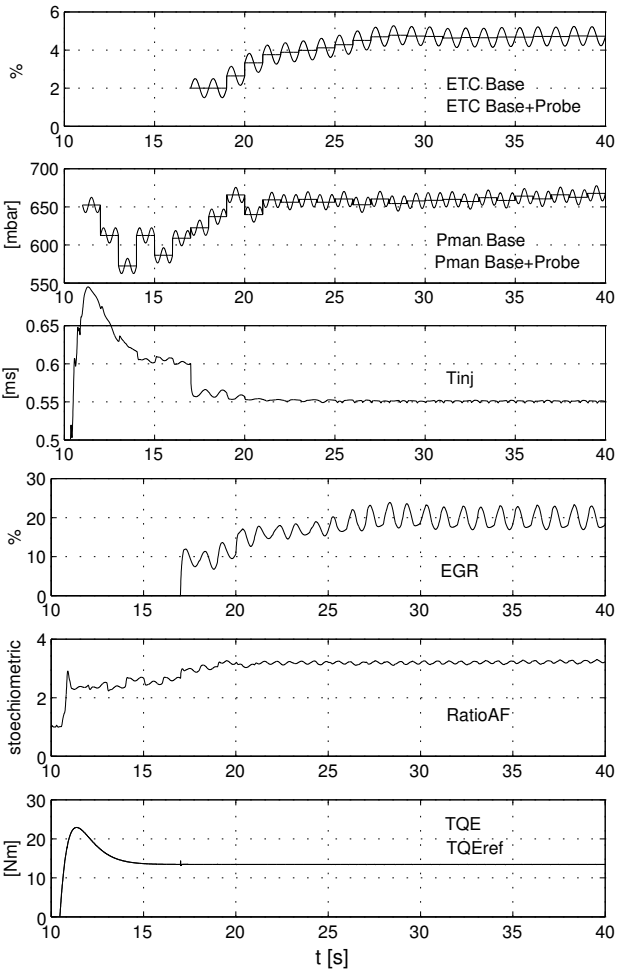

Figure 14 Results from driving-cycle scenario at gear 3 and velocity 40 km/h with extremum control of $R_{\mathrm{a/f}}$ and EGR ($N \approx 1900$ rpm). Note how P_{man} and ETC simultaneously converge to optimal base values.

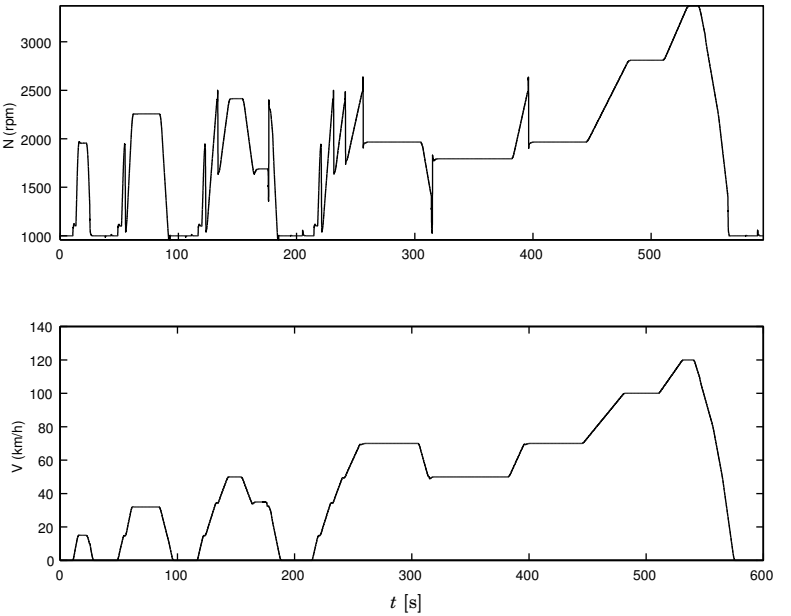

Figure 15 Engine speed and vehicle speed with the complete engine management system in a simulation of the European driving cycle. Gear switches and idle-speed control engagement and disengagement can be seen in the engine speed signal.

geneous to stratified mode takes place immediately before the gear-switch. After the gear-switch the system switches back to homogeneous mode. Small irregularities in the torque signal occur during the mode-changes. During a short interval the idle-speed controller becomes active.

The controller was evaluated with the same configurations as for the core controller. The simulations are evaluated using three different criteria values that were defined in the benchmark:

- $C_{\text{Throttle}} = T_{\text{samp}} \cdot \sum |MTC(k) - MTC(k-1)| / T_{\text{scenario}}$ is a measure of the variations in MTC_{pos}.
- $C_{\text{Cons}} = \int Q_{\text{fuel}} dt / \left(\rho_{fuel} \cdot \int v dt\right)$ is a measure of the fuel consumption (liters / 100 kilometers)
- C_{Global} is the average value of five criteria that measure the performance of the air-fuel ratio control in homogeneous mode, penalize large torque-gradients, penalize torque bumps caused by mode changes, measure the performance of the vehicle-speed control, and measure the performance of the idle-speed control.

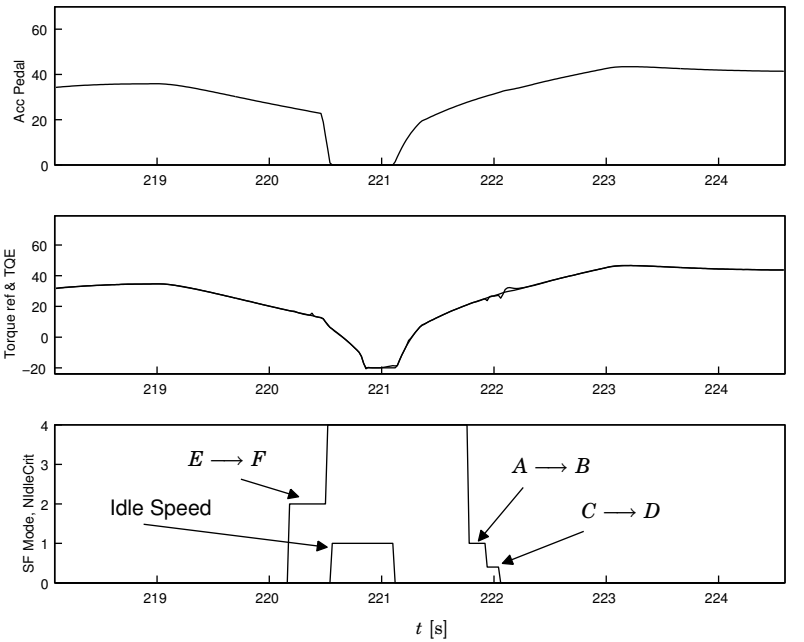

Figure 16 Accelerator pedal (top), torque and torque reference (middle), and combustion mode and idle-speed mode during a gear-switch phase. The mode-switching states as defined above are indicated in the figure.

All three criteria should be minimized.

During the full driving-cycle 26 combustion-mode changes are performed. The engine spends about the same amount of time in homogeneous mode as in stratified mode. The perturbation signal of the extremum controller was chosen as $\delta P_{\text{man}} = 5\sin(2.5\pi t)$ mbar, which is significantly smaller and slower than in the core controller simulations. This is to reduce the cost of throttle actuation that is part of the evaluation criteria. The mean fuel consumption was 6.78 l/100km, compared to 6.88 l/100km for a controller running with fully-open throttle in stratified mode. The results of the present work thus indicate that throttling may be used to improve fuel consumption. It should be emphasized that these results apply to the benchmark engine model, and that it is uncertain how they translate to a real engine.

The controller has been compared with two other designs: *(MPC)*, a model-predictive controller based on a piecewise-linear model of the GDI engine, developed using clustering-based identification on the Simulink engine model [Mollov *et al.*, 2001], and *(PWL)*, a design based on feedfor-

ward using piecewise-linear inverse-approximations of the different engine-maps, combined with feedback [Passaquay, 2000]. The latter control design is relatively close in spirit to a production engine management system, with the exception that piecewise-linear models are used instead of look-up tables. A summary of the results is shown in Figure 17. The approach taken in the present work proved to yield better fuel economy and tracking performance for all configurations. The throttle command-signal cost C_{throttle} is somewhat higher in configurations 1–3, due to the probing actions of the extremum controller. This cost may be reduced by a more clever choice of probing signals. There is no significant deterioration in throttle cost in the noisy configurations 4–6. The better result with respect to fuel consumption can be explained by the online optimization using the extremum controller, in combination with the torque control algorithms. In the PWL design ignition timing is used for fast torque control, which results in lower efficiency of the engine, and as a result higher fuel consumption. In the global criteria, which is essentially a measure of tracking performance and the ability to switch mode without large torque bumps, it is interesting to see that the present design is significantly more robust to gain variations from the nominal design in configuration 1. This is due to the extensive use of feedback, which results in robustness to engine variations. Also here the introduction of noise does not significantly deteriorate performance.

6. Conclusions

Conventional engine management systems are dominated by feedforward control based on look-up tables containing inverse approximations of different engine maps. The reason for this is the lack of inexpensive sensors for, e.g., engine torque and in-cylinder pressure. Such open-loop approaches are sensitive to engine-to-engine variations and aging and wear phenomena. Since feedback is not possible without a sensor, the look-up tables must be very accurate. Development of the tables is time-intensive and costly. This is one of the reason for the large interest in the engine-control community for different data-based nonlinear approximation schemes, e.g., neural networks and fuzzy systems.

The work presented in this paper indicates that much simpler control designs may be used with very good results, provided that the effective torque produced by the engine is available for feedback. The work has aimed at the design of a robust, simple and efficient EMS for a GDI engine. A bottom-up approach has been adopted, where standard building blocks like linear PID-controllers and static affine feedforward have been used extensively for building a controller for torque, idle-speed, and

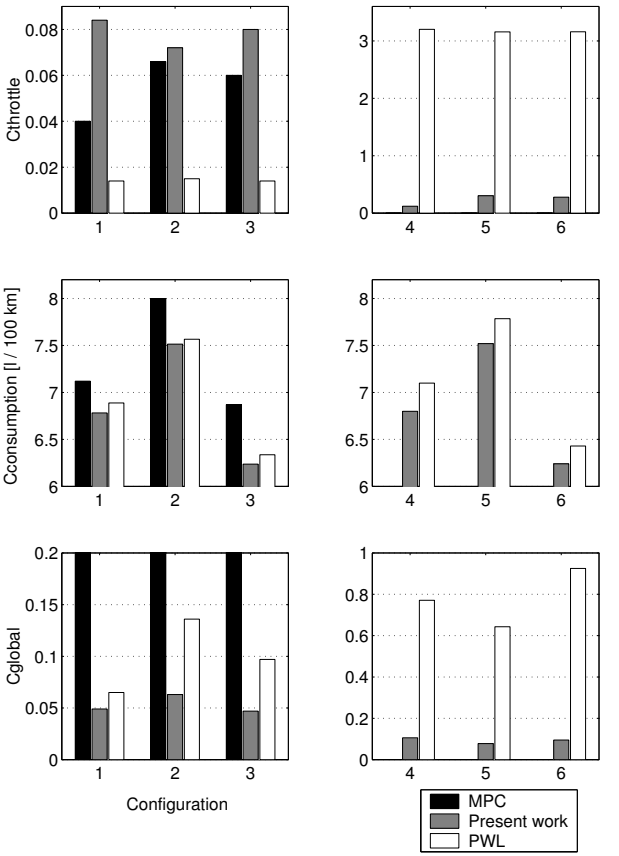

Figure 17 Criteria results. The MPC design results for C_{global} has been truncated to keep the results from the present work and the PWL design distinguishable. There are no results available for the MPC design for configurations 4–6.

combustion-mode control. The controller also includes extremum control of air-fuel ratio and EGR to minimize fuel consumption and emissions. The feedback controllers rely on the availability of the effective torque signal. The silent extremum control strategy that is developed also requires torque feedback to eliminate effects of perturbation signals on the engine output. The suggested silent extremum-control structure may find applications in other similar control problems. It has the useful property of optimizing a target function with respect to a control variable by adding disturbances that are invisible in the system output.

The EMS was evaluated in computer simulations on the full European

driving cycle with good results that compare well with traditional table based engine control with respect to tracking performance, ride comfort, and fuel consumption. The strategy was proven to be robust to engine variations and disturbances.

The results show that it is possible to design engine management systems with less prior knowledge of engine model parameters if feedback control and online optimization are used extensively. The reasons why the control design works as well as it does are the inherent robustness of feedback to uncertainties, the linearizing nature of feedback, the extremum-based model-free on-line optimization, and the switching strategy that minimizes fuel consumption and torque bumps.

Acknowledgements

This work has been performed within the EU Esprit FAMIMO project. The work of the first author has been supported by the DICOSMOS project within the NUTEK Complex Technical Systems Program, and the work of the third author has been supported financially by The Danish Steel Works Ltd. We would like to thank Dominique Passaquay, Pascal Bortolet, and Serge Boverie at Siemens Automotive, Toulouse, and Bengt Johansson at Department of Heat and Power Engineering, Division of Combustion Engines, Lund Institute of Technology.

References

Andersson, P., L. Eriksson, and L. Nielsen (1999): "Modeling and architecture examples of model based engine control." In *Proceedings of the Second Conference on Computer Science and Systems Engineering*. Linköping, Sweden.

Åström, K. J. and B. Wittenmark (1997): *Computer-Controlled Systems*, third edition. Information and System Sciences Series. Prentice-Hall.

Azzoni, P., D. Moro, F. Ponti, and G. Rizzoni (1998): "Engine and load torque estimation with application to electronic throttle control." SAE Technical Paper 980795.

Draper, C. and Y. Li (1951): "Principles of optimalizing control systems and an application to the internal combustions engine." *ASME*.

Druzhinina, M., I. Kolmanovsky, and J. Sun (1999): "Hybrid control of a gasoline direct injection engine." In *Proceedings of the 38th conference on Decision & Control*, pp. pp. 2667–2672. Phoenix, Arizona USA.

Fleming, W. J. (1989): "Magnetostrictive torque sensor performance — nonlinear analysis." *IEEE Transactions on Vehicular Technology*, **38:3**, pp. 159–167.

Fons, M., M. Muller, A. Chevalier, C. Vigild, E. Hendricks, and S. C. Sorenson (1999): "Mean value engine modelling of an SI engine with EGR." SAE Technical Paper 1999-01-0909.

Gäfvert, M., K.-E. Årzén, and L. M. Pedersen (2000a): "Simple linear feedback and extremum seeking control of GDI engines." In *Proceedings of Seoul 2000 FISITA World Automotive Congress*. Seoul, Korea.

Gäfvert, M., L. M. Pedersen, K.-E. Årzén, and B. Bernhardsson (2000b): "Simple feedback control and mode switching strategies for GDI engines." In *SAE 2000 World Congress*. Detroit, Michigan, USA. SAE Technical Paper 2000-01-0263.

Gerhardt, J., H. Honninger, and H. Bischof (1998): "A new approach to functional and software structure for engine management systems." SAE Technical Paper 980801.

Gyan, P., S. Ginoux, J.-C. Champoussin, and Y. Guezennec (2000): "Crankangle based torque estimation: Mechanistic/stochastic." SAE Technical Paper 2000-01-0559.

Haskara, I. and L. Mianzo (2001): "Real-time cylinder pressure and indicated torque estimation via second order sliding modes." In *Proceedings of the American Control Conference*, pp. pp. 3324–3328. Arlington.

Hendricks, E. (1990): "Mean value modelling of spark ignition engines." SAE Technical Paper 900616.

Hendricks, E. (1997): "Engine modelling for control applications: A critical survey." *Meccanica*, **32**, pp. pp. 387–396.

Hendricks, E., D. Engler, and M. Fam (2000): "A generic mean value engine model for spark ignition engines." In *41st Simulation Conference, SIMS 2000*. Technical University of Denmark, Lyngby, Denmark.

Hendricks, E., M. Jensen, A. Chevalier, and T. Vesterholm (1994): "Problems in event based engine control." In *Proccedings of the American Control Conference*, pp. pp. 1585–1587. Baltimore, Maryland.

Iwamoto, Y., K. Noma, O. Nakayama, T. Yamauchi, and H. Ando (1997): "Development of gasoline direct injection engine." SAE Technical Paper 970541.

Jackson, N. S., J. Stokes, P. A. Whitaker, and T. H. Lake (1997): "Stratified and homogeneous charge operation for the direct injection gasoline engine — high power with low fuel consumption and emissions." SAE Technical Paper 970543.

Jeremiasson, J. and C. Wallin (1998): "Balancing of individual cylinders in a V8 diesel engine based on crankshaft torque measurements." SAE Technical Paper 981063.

Johansson, R. and A. Rantzer, Eds. (2002): *Nonlinear and Hybrid Systems in Automotive Control*, chapter 13. Springer-Verlag London.

Kiencke, U. (1987): "The role of automatic control in automotive systems." In *Proceedings of the IFAC 10th Triennal World Congress*. Munich, FRG.

Kolmanovsky, I., M. van Nieuwstadt, and J. Sun (1999): "Optimization of complex powertrain systems for fuel economy and emission." In *Proceedings of the 1999 IEEE International Conference on Control Applications*, pp. pp. 833–839. Hawaii, USA.

Kume, T., Y. Iwamoto, K. Lida, M. Murakami, K. Akishino, and H. Ando (1996): "Combustion control technologies for direct injection SI engine." SAE Technical Paper 960600.

Mollov, S., R. Babuska, and P. J. van der Veen (2001): "Nonlinear model based predictive control of a GDI engine." *Journal A*, **42:1**, pp. 31–38.

Moskwa, J. J. and C.-H. Pan (1995): "Engine load torque estimation using nonlinear observers." In *Proceedings of the 34th Conference on Decision & Control*, pp. pp. 3397–3402. New Orleans, LA, USA.

Passaquay, D. (2000): *Modélisation et Commande de processus multivariables à base de logique floue: Application à la commande de moteurs thermiques.* PhD thesis, Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS. (Referred sections in English.).

Passaquay, D., P. Bortolet, and S. Boverie (1998): "Engine benchmark architecture decomposition, reduced benchmark, controller design rules." Technical Report. Siemens, Toulouse. Release 1.2 (Confidential).

Rizzoni, G. (1989): "Estimate of indicated torque from crankshaft speed fluctuations: A model for the dynamics of the IC engine." *IEEE Transactions on Vehicular Technology*, **38:3**, pp. pp. 168–179.

Sasaki, S., D. Sawada, T. Udeda, and H. Sami (1998): "Effects of EGR on direct injection gasoline engine." *JSAE Review*, **No 19**, pp. 223 – 228.

Sobel, J. R., J. Jeremiasson, and C. Wallin (1996): "Instantaneous crankshaft torque measurements in cars." SAE Technical Paper 960040.

Sternby, J. (1979): "A review of extremum control." Technical Report LUTFD2/(TFRT–7161)/1-47/(1979). Department of Automatic Control, Lund Institute of Technology.

Sun, J., I. Kolmanovsky, D. Brehob, J. A. Cook, J. Buckland, and M. Haghgooie (1999): "Modeling and control of gasoline direct injection stratified charge (DISC) engines." In *Proceedings of the 1999 IEEE International Conference on Control Applications*, pp. 471–477.

Wallin, C., L. Gustavsson, and M. Donovan (2002): "Engine monitoring of a formula 1 racing car based on direct torque measurements." SAE Technical Paper 02P-195.

Wang, Y.-Y., V. Krishnaswami, and G. Rizzoni (1997): "Event-based estimation of indicated torque for IC engines using sliding-mode observers." *Control Engineering Practice*, **5:8**, pp. pp. 1123–1129.

Zao, F.-Q., M.-C. Lai, and D. L. Harrington (1997): "A review of mixture preparation and combustion control strategies for spark-ignited direct-injection gasoline engines." SAE Technical Paper 970627.

Paper IV

On the Effect of Transient Data-Errors in Controller Implementations

Magnus Gäfvert Björn Wittenmark Örjan Askerdal

Abstract

Computer-level faults leading to data errors in computations are predicted to occur increasingly frequent in future microprocessors. This work discusses the impact of such errors on closed-loop performance in implementations of digital control systems. A method to render a control system more robust to data errors by introducing artificial signal limits and then combine them with an anti-windup scheme is presented and exemplified.

1. Introduction

As computers, rather than electro-mechanical systems, are increasingly used for implementing control algorithms, control systems become more vulnerable to computer level failures due to faults in semiconductor devices. Such faults are classified as: (i) *transient faults* which are short-duration faults that are induced by neutron and alpha particles, power supply and interconnect noise, and electrostatic discharge, (ii) *intermittent faults* which are re-occurring short-term faults that occur due to marginal hardware or aging effects, and (iii) *permanent faults* which have the same causes as the intermittent faults, but reflect irreversible physical changes. The trend with decreasing transistor and interconnect dimensions, lower power supply, and higher operating frequencies contributes to increasing the occurrence rates of transient and intermittent faults, while improvements in design and manufacturing have led to decreasing permanent failure rates, less than 15 FIT (Failures In Time, or failures in 10^9 hours of operation) for microprocessors in 2001 [Constantinescu, 2002]. This paper will therefore focus on transient faults.

A microprocessor is built from SRAM memory cells, latches, and combinational logic. Faults in the microprocessor may either result in control-flow errors, in which case the instruction execution order is erroneous, or in data errors, in which case the executing program delivers erroneous results. Many control-flow errors may be detected with watch-dog timers. Modern microprocessors have built-in protection against transients faults in memory cells (cache) using e.g. error-correcting codes and parity checks, while the remaining parts are essentially unprotected. Methods to implement protection against transient faults in other locations than memory cells often involve redundant micro-controller subsystems (ALU etc.) as in [Slegel *et al.*, 1999], but they result in more complex and more expensive devices. Hence, remaining failures are data errors due to transient faults in the combinational logic or latches. In [Shivakumar *et al.*, 2002] it is predicted that the transient fault rate per chip of combinational logic circuits will increase nine orders of magnitude from 1992 ($\sim 10^{-7}$ FIT) to 2011 ($\sim 10^2$ FIT), when the failure rate will be comparable to that of unprotected memory elements. The corresponding fault rate for the present generation of microprocessors is ~ 1 FIT. The trend of increasing fault rates is also reported in [Hazucha, 2000], where it is predicted that a 32 Mbit static memory implemented in a 0.1 μm process will fail on the average each 5.7 years at sea level. A result that is expected to hold also for other logic circuits, such as flip-flops, latches, registers, and combinational circuits.

In safety-critical applications it is imperative to have error detection and recovery functionality to meet the high demands on dependability. In

some areas, such as the aircraft industry, fault-tolerance is achieved by redundant hardware and high-end devices. In more cost-sensitive areas, such as the automotive industry, expensive solutions may not be feasible. Assuming a microprocessor with a (constant) transient failure rate of 10 FIT that hosts e.g. a braking functionality in a car, the mean time to computer failure (possibly leading to catastrophic failure for the vehicle) is more than 11,415 years for a single vehicle. For a series of 100,000 cars, the mean time to failure in one of the cars is 42 days. If a microprocessor executes a control algorithm that is likely to tolerate a transient data error, this device may not necessarily have to be as fault-tolerant, and can thus be less expensive. Therefore it is of interest to investigate the impact of transient data errors on hardware hosting implementations of safety-critical control algorithms.

Understanding the effect of data errors on general computer functionality is an intensively researched area, e.g., [Pradhan, 1996]. Methods for analyzing the effects of timing errors on control systems were presented in, e.g., [Kim and Shin, 1994, Wittenmark *et al.*, 1995]. Analysis of effects on system stability of data errors caused by EMI bursts was investigated in [Kim *et al.*, 2000]. However, catastrophic failures in a safety-critical system may occur before the system reaches instability, e.g. if some constraint on the control error is exceeded. Recent results [Cunha *et al.*, 2001, Vinter, 2001] show that many data errors will have a limited effect on control performance, i.e., control systems often have an inherent resilience or inertia to data errors. This path is pursued further in this paper. It is discussed how the implementation of the control algorithms influences the tolerance to data errors resulting from computer node failures. Related work is published in [Askerdal *et al.*, 2002], where classical frequency-domain methods are used to investigate the effects of computer failures on linear feedback control systems.

The paper is organized as follows: In Section 2 a general model for the effect of data errors on a control algorithm is stated. Thereafter, controller realization and scaling is discussed in Section 3. Sections 4 and 5 show how computed signal bounds together with an anti-windup scheme may improve the recovery from data errors. This is illustrated with an example in Section 6. The results are discussed in Section 7, and summarized in Section 8.

2. Modeling Data Errors Caused by Computer Node Faults

Let

$$\begin{aligned} z(k+1) &= \Phi_c z(k) + \Gamma_c^{u_c} u_c(k) + \Gamma_c^{y} y(k) \\ u(k) &= C_c z(k) + D_c^{u_c} u_c(k) + D_c^{y} y(k) \end{aligned} \tag{1}$$

be a state-space realization of a general linear two-degrees-of-freedom dicrete-time controller with internal state z, command signal u_c, process output y, and control signal u. The controlled process is assumed to be a linear time-invariant system

$$Y(s) = G(s)U(s) \tag{2}$$

with zero-order-hold sampling. A pseudo-code implementation of (1) would be

```
repeat:
   uc := read_from_interface();
   y := read_from_sensors();
   u := compute_control_signal(z,uc,y);
   write_to_actuator(u);
   z := update_controller_state(z,uc,y);
   wait_for_next_sample;
```

This algorithm would be implemented and executed in a computer node like in Figure 1. Erroneous computation results due to transient data errors would eventually be stored in u or z, and would propagate to the controlled process, and through the feedback loop. As illustrated in the figure, the internal components of a computer node are: communication controllers, memories, microprocessors, and internal communication buses. All these components may be affected by faults. As the communication

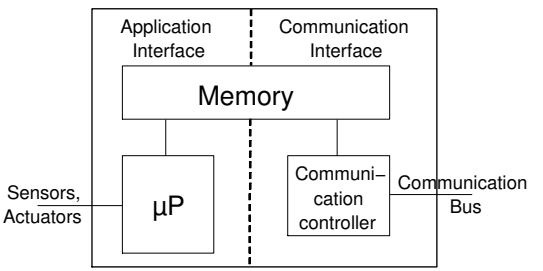

Figure 1 A General architecture of a computer node.

controller handles the information exchange with other nodes, quantities measured by sensors may become erroneous due to faults in the communication controller. The data path of a microprocessor consists of caches, registers, buses, and functional units (i.e., ALU, multiplier, etc.). If a fault occurs in any of these parts, it will affect the ongoing calculation if the faulty part is activated, rendering the result of that calculation to be incorrect. The communication inside the computer node is, generally, performed using buses. If a fault occurs on such a bus, the data currently

being transferred will be affected and the result may be a data error in any of the calculated data.

Data errors may be regarded as bit-flips in the digital representation of the affected variable or signal. Single-bit errors tend to be more common than multiple-bit errors [Askerdal, 2000]. In a N-bit fixed-point representation with M fractional bits the number range is $R = [-2^{N-M-1}, 2^{N-M-1} - 2^{-M}]$, with a resolution of $Q = 2^{-M}$. With this representation, the magnitudes of the errors will be in the same range as the control signals (assuming that the data have been properly scaled). In the IEEE floating-point standard with f fraction bits, and e exponent bits the range is $R = [-(2-2^{-f}) \cdot 2^{2^{e-1}-1}, (2-2^{-f}) \cdot 2^{2^{e-1}-1}]$, and bit flips in the most significant exponent bits may lead to very large errors. A reasonable model of data errors is additive impulse disturbances with magnitude of rectangular distribution within the number representation range. The data error disturbances may then be included in (1) as

$$\begin{aligned} z(k+1) &= \Phi_c z(k) + \Gamma_c^{u_c} u_c(k) + \Gamma_c^{y} y(k) + \eta_z(k) \\ u(k) &= C_c z(k) + D_c^{u_c} u_c(k) + D_c^{y} y(k) + \eta_u(k) \end{aligned} \tag{3}$$

where $\eta_z(k)$ and $\eta_u(k)$ are impulse disturbances due to data errors affecting the computation of the controller state and control signal, respectively. Since transient data errors occur sporadically, it may be assumed that an error manifests in either $\eta_z(k)$ or $\eta_u(k)$.

3. Controller Realizations

With a change of state variables $\bar{z}(k) = Tz(k)$ in (3) the controller realization becomes

$$\begin{aligned} \bar{z}(k+1) &= \bar{\Phi}_c \bar{z}(k) + \bar{\Gamma}_c^{u_c} u_c(k) + \bar{\Gamma}_c^{y} y(k) + \eta_z(k) \\ u(k) &= \bar{C}_c \bar{z}(k) + D_c^{u_c} u_c(k) + D_c^{y} y(k) + \eta_u(k) \end{aligned} \tag{4}$$

with $\bar{\Phi}_c = T\Phi_c T^{-1}$, $\bar{\Gamma}_c^{u_c} = T\Gamma_c^{u_c}$, $\bar{\Gamma}_c^{y} = T\Gamma_c^{y}$ and $\bar{C}_c = C_c T^{-1}$. Note that η_z is not transformed in (4), since it appears internally in the controller implementation. Combining the controller realization (4) with a state space realization of the sampled controlled process (2)

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Gamma u(k) \\ y(k) &= Cx(k) \end{aligned} \tag{5}$$

results in the closed-loop dynamics

$$\begin{aligned} \xi(k+1) &= \Phi_{cl}\xi(k) + \Gamma_{cl}^{u_c}u_c(k) + \Gamma_{cl}^{\eta_z}\eta_z(k) + \Gamma_{cl}^{\eta_u}\eta_u(k) \\ y(k) &= C_{cl}^{y}\xi(k) \\ u(k) &= C_{cl}^{u_c}\xi(k) + D_c^{u_c}u_c(k) + \eta_u(k) \end{aligned} \tag{6}$$

with $\xi = (x, \bar{z})$ and

$$\Phi_{cl} = \begin{pmatrix} \Phi + \Gamma D_c^y C & \Gamma C_c T^{-1} \\ T\Gamma_c^y C & T\Phi_c T^{-1} \end{pmatrix}$$

$$\Gamma_{cl}^{u_c} = \begin{pmatrix} \Gamma D_c^{u_c} \\ T\Gamma_c^{u_c} \end{pmatrix} \quad \Gamma_{cl}^{\eta_z} = \begin{pmatrix} 0 \\ I \end{pmatrix} \quad \Gamma_{cl}^{\eta_u} = \begin{pmatrix} \Gamma \\ 0 \end{pmatrix}$$

$$C_{cl}^{y} = \begin{pmatrix} C & 0 \end{pmatrix} \quad C_{cl}^{u} = \begin{pmatrix} D_c^y C & C_c T^{-1} \end{pmatrix}$$

The impulse-responses from $\eta_z(k)$ and $\eta_u(k)$ to $y(k)$ are $h_{\eta_z}(k) = C_{cl}^{y}\Phi_{cl}^{k-1}\Gamma_{cl}^{\eta_z}$, $k > 0$, and $h_{\eta_u}(k) = C_{cl}^{y}\Phi_{cl}^{k-1}\Gamma_{cl}^{\eta_u}$, $k > 0$. Inspection of Φ_{cl} reveals that the structure

$$\Phi_{cl}^{k-1} = \begin{pmatrix} * & *T^{-1} \\ T* & T*T^{-1} \end{pmatrix} \tag{7}$$

is preserved for all $k > 1$. Hence, the impulse-responses will have the structures $h_{\eta_z}(k) = C*T^{-1}$ and $h_{\eta_u}(k) = C*\Gamma$ for $k > 1$. Obviously $h_{\eta_u}(k)$ is invariant under the change of state variables, while $h_{\eta_z}(k)$ is not. With a large diagonal state scaling $T = \Lambda$, the output response to impulse errors on the states can be made arbitrary small. However, the state scaling is constrained by the available numeric precision.

The problem to find the controller realization, i.e. find the T, such that the output response $y(k)$ to the disturbance $\eta_z(k)$ is minimized with respect to some measure, under the constraint that the state scaling is kept appropriate for the available numeric representation is treated in e.g. [Moroney *et al.*, 1983, Rotea and Williamson, 1995]. The results are optimal realizations in the context of round-off errors, which have the same problem structure, but where the disturbances are close to stationary white noise processes. With $y(k) = H_{\eta_z}(q)\eta_z(k)$, $\bar{z}(k) = H_z(q)u_c(k)$, solutions to the minimization of $\|H_{\eta_z}\|_p$ with respect to T, subject to the state scaling constraint $\|H_z\|_q \leq \gamma$ are presented for $p, q \in \{2, \infty\}$ in [Rotea and Williamson, 1995]. For data errors it may be more natural to regard $p = q = 1$, i.e. peak-to-peak gain. Even more natural would it be to minimize $\|h_{\eta_z}\|_\infty$, since the disturbances are assumed to be impulses. However, in practice it is likely more feasible to optimize the realization with respect to the all-time present round-off noise than rare and sporadic data errors. In the following it will thus be assumed that the transformation T is

given, and we concentrate on the analysis of the influence of η_z and η_u on the closed-loop performance. Hence, we can consider the representations (3) and (5) of the controller and the process.

4. Signal Bounds

If a bound on the command signal u_c is known, then l_∞ bounds on controller states and the control signal may be computed using l_1 norms on impulse responses [Hanselmann, 1987]. The controller state and control signal are given by

$$z_i(k) = \sum_{j=0}^{k} h_i(k-j)u_c(j) \tag{8a}$$

$$u(k) = \sum_{j=0}^{k} h_u(k-j)u_c(j) \tag{8b}$$

where $\{h_i(k)\}$ and $\{h_u(k)\}$ are the impulse-response sequences of the closed-loop system from the command signal $u_c(k)$ to the state variable $z_i(k)$ and control signal $u(k)$. Bounds on z_i may be computed by applying the Hölder inequality on (8a):

$$|z_i(k)| \le \sum_{j=0}^{k} |h_i(k-j)u_c(j)| \le \left[\sum_{j=0}^{k} |h_i(j)|^p\right]^{1/p} \left[\sum_{j=0}^{k} |u_c(j)|^q\right]^{1/q} \tag{9}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. If the maximum absolute value of the command signal is known $|u_c(k)| \le M$, then Equation (9), with $p = 1$ and $q = \infty$, gives

$$|z_i(k)| \le M \sum_{j=0}^{k} |h_i(j)| \tag{10a}$$

Correspondingly a bound on $u(k)$ may be expressed as

$$|u(k)| \le M \sum_{j=0}^{k} |h_u(j)| \tag{10b}$$

It is straightforward to include the effects of other bounded inputs to the closed-loop system, e.g. load disturbances, in the bounds of Equations (10). The bound (10a) may also be used for l_1 state scaling, as described in [Hanselmann, 1987]. If the closed-loop system is well designed with proper damping, the bounds are not expected to be very conservative, which is

also confirmed in [Hanselmann, 1987]. If the signal bounds are exceeded it can be concluded that an error has occurred in the system. Hence, the signal bounds may be used on-line to actively detect deviations from normal operation, similarly to the approach in e.g. [Stroph and Clarke, 1998]. Note, however, that in the cited work the bounds appear to be computed on the controller in open-loop, which result in over-estimation of the bounds by factors of magnitudes in comparison with the closed-loop bounds. The approach taken in the present work is to introduce explicit bounds in the controller, that correspond to those of (10), and then use well-known anti-windup methods to handle these signal limitations in a graceful manner. This will give the system an inherent robustness to data errors that exceed the signal bounds.

5. Anti-windup

In the presence of control signal limitations, the control signal actually delivered to the controlled process will be $u(k) = \text{sat}(v(k))$, where $v(k)$ is the linear control signal. When the output signal is saturated the feedback path is broken and the controller states are driven in open-loop, leading to deteriorated performance or even instability. If the controller has integral action this phenomenon is denoted integrator windup. To inhibit this behavior various anti-windup schemes may be applied [Edwards and Postlethwaite, 1998]. Anti-windup should always be implemented in a controller with actuator saturation. In this work an explicit artificial limitation according to (10b) is introduced to make the system robust to data errors. In practice an actuator limitation will also be present. If the actuator saturation limit is smaller than the artificial limit (10b), then the smallest limit should be used. The observer-based anti-windup of [Åström and Wittenmark, 1997] is a general method where the control signal error is fed back to the controller. With observer-based anti-windup the controller (3) is modified as

$$\begin{aligned} z(k+1) &= \Phi_c z(k) + \Gamma_c^y y(k) + \Gamma_c^{u_c} u_c(k) \\ &\quad + K\left(u(k) - v(k)\right) + \eta_z(k) \\ &= \hat{\Phi}_c z(k) + \hat{\Gamma}_c^y y(k) + \hat{\Gamma}_c^{u_c} u_c(k) \\ &\quad + Ku(k) + \eta_z(k) \\ v(k) &= C_c z(k) + D_c^y y(k) + D_c^{u_c} u_c(k) + \eta_u(k) \\ u(k) &= \text{sat}(v(k)) \end{aligned} \tag{11}$$

with $\hat{\Phi}_c = \Phi_c - KC_c$, $\hat{\Gamma}_c^y = \Gamma_c^y - KD_c^y$, and $\hat{\Gamma}_c^{u_c} = \Gamma_c^{u_c} - KD_c^{u_c}$. The gain K is chosen as to obtain the desired observer dynamics given by $\hat{\Phi}_c$. The anti-

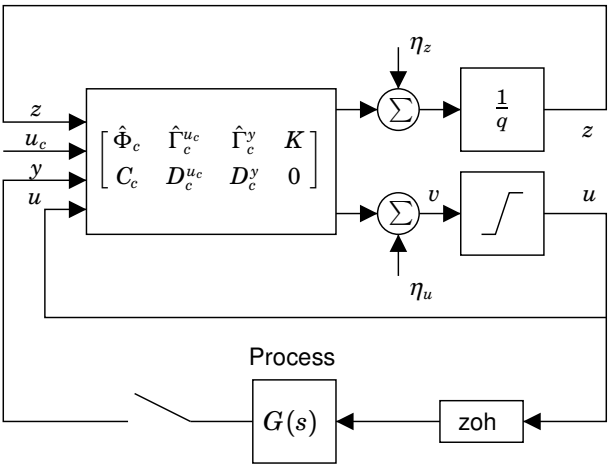

Figure 2 Blockdiagram of closed-loop system with artificial control-signal bound and anti-windup.

windup scheme will now reduce the effect of data errors that are causing the controller output to exceed the estimated limits. Note that the observer based anti-windup operates on all controller states, while certain other schemes only operates on the integrator state. Hence, errors in any of the controller states are eliminated due to the anti-windup. Also note that the observer-based method does not require any additional states to be introduced. In the context of data errors this is important, since data errors affecting explicit anti-windup states would not be handled gracefully by the system. The closed-loop system resulting from combining (11) with (2) is shown in Figure 2.

6. Example

As an illustration of the inherent tolerance to transient faults that may be achieved with the combination of signal bounds and anti-windup, we study the control of the simple servo process

$$G(s) = \frac{100}{s(s+10)} \tag{12}$$

A discrete-time two-degrees-of-freedom tracking controller is synthesized using the polynomial pole-placement design method of [Åström and Wittenmark, 1997]. The controller is designed for a closed-loop bandwidth of

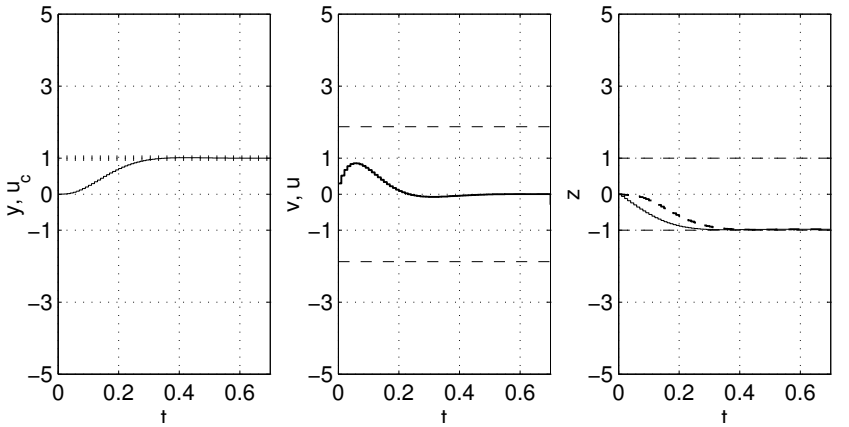

Figure 3 Closed-loop system response to a command signal step. Computed bounds on u and z are shown. (Left: u_c (dotted), y (solid); Middle: u (solid), bounds (dashed); Right: z_1 (solid), z_2 (dashed), bounds (dashed))

$\omega_{cl} = 15$ rad/s, observer dynamics of $2\omega_{cl}$, and integral action. The sampling time is set to 0.01 s. A minimal second-order modal-form state-space realization of the controller is used for implementation, with $z = (z_1, z_2)$. In this realization z_1 is an integrator state, and z_2 may be interpreted as an observer state. The command signal u_c is assumed to be bounded by $|u_c(k)| \leq 1$. The l_1 state-scaling of (10a) is applied such that $|x_i(k)| \leq 1$. The control signal bound according to (10b) is $|u(k)| \leq 1.88$. (The corresponding open-loop bound according to [Stroph and Clarke, 1998] is $|u(k)| \leq 341$.) The anti-windup scheme of (11) is implemented, with K chosen as to obtain dead-beat dynamics, to obtain the closed-loop system of Figure 2. In Figure 3 the closed-loop step-response to a command-signal unit step is shown. Note that the computed state-bounds are very tight, while the control signal bound seems to be a little more conservative. In Figure 4 the closed-loop response to a data error $\eta_{z_1}(k) = 5\delta(k)$, affecting the computation of the integrator state z_1, is shown in the case when the artificial signal-limit and anti-windup are not applied. It can be seen how the integrator state slowly recovers, while the control error grows large. In presence of actuator limitation large data-error amplitudes even result in instability. Figure 5 shows the corresponding response with application of the artificial signal limit $|u(k)| \leq 1.88$, in combination with dead-beat anti-windup. In this case the controller state recovers within a few samples, and the control error is moderate. Figure 6 shows the largest absolute control-error that results from various magnitudes on the data errors η_u, η_{z_1}, and η_{z_2}. Four different system settings are shown for comparison: (i)

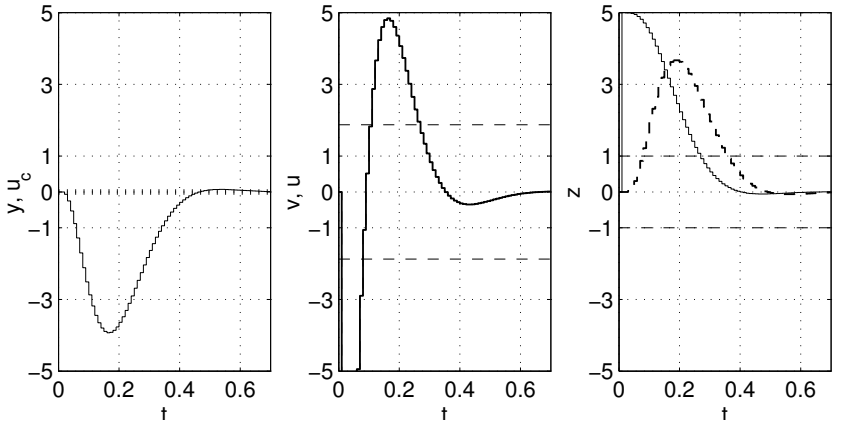

Figure 4 Closed-loop system response to a data error $\eta_{z_1}(k) = 5\delta(k)$, on controller state z_1, without signal limitation and anti-windup. Computed bounds on u and z are shown. (Left: u_c (dotted), y (solid); Middle: u (solid), bounds (dashed); Right: z_1 (solid), z_2 (dashed), bounds (dashed))

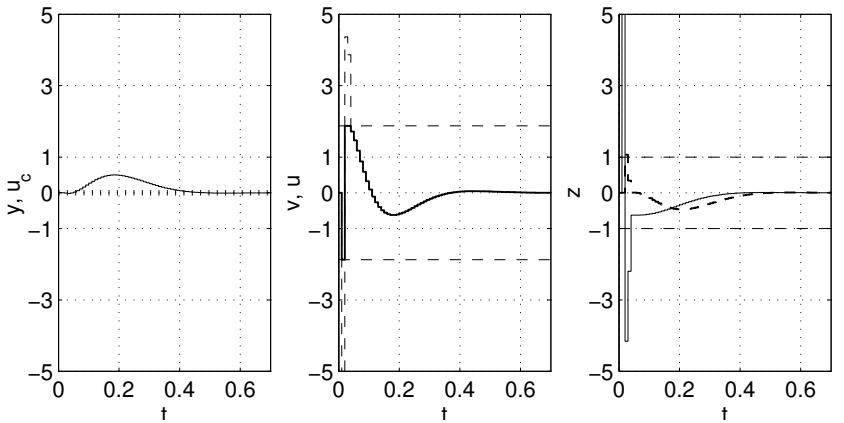

Figure 5 Closed-loop system response to a data error $\eta_{z_1}(k) = 5\delta(k)$, on controller state z_1, with signal limitation and dead-beat anti-windup. Computed bounds on u and z are shown. (Left: u_c (dotted), y (solid); Middle: u (solid), v (dashed), limits (dashed); Right: z_1 (solid), z_2 (dashed), bounds (dashed))

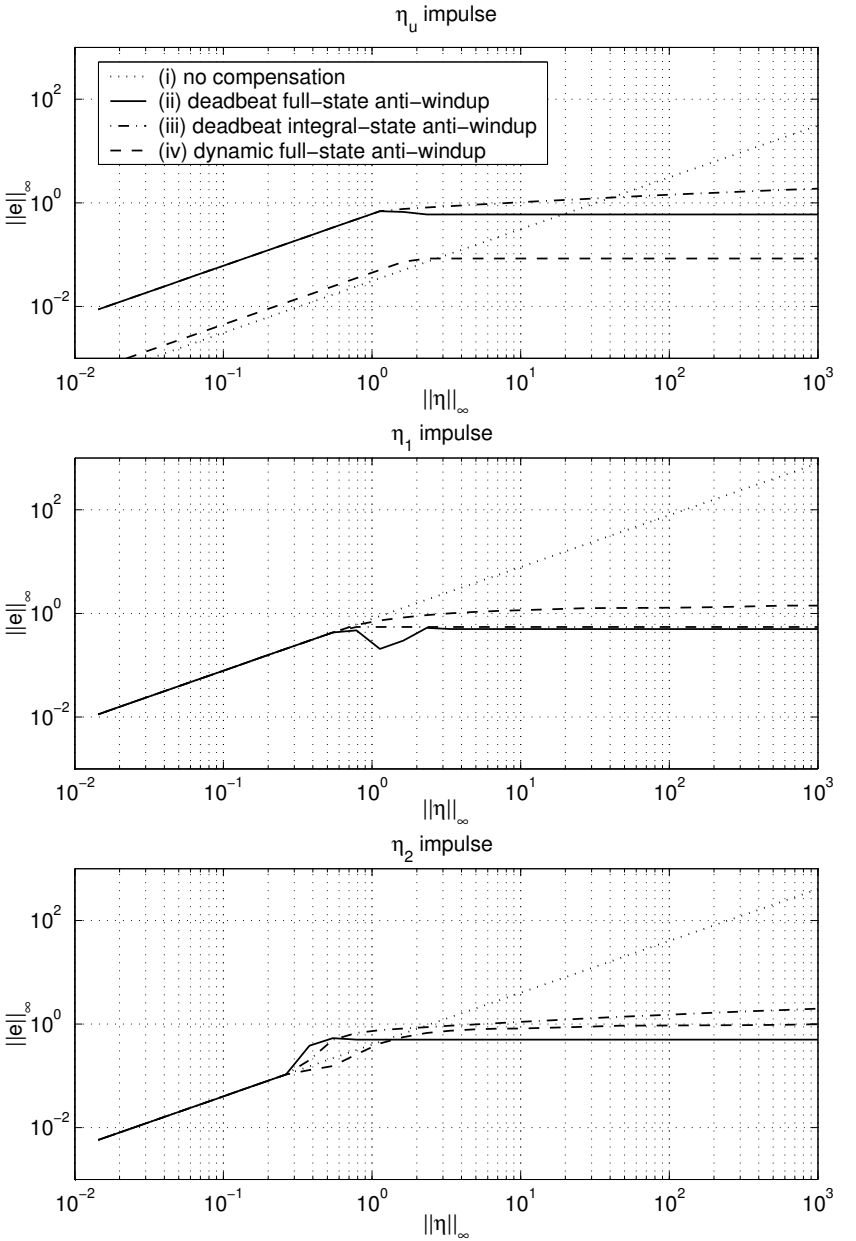

Figure 6 Closed-loop system response to impulse data errors. Control-error magnitude versus data-error magnitude. (Upper: η_u; Middle: η_{z_1}; Lower: η_{z_2})

No control-signal limits or anti-windup. Here the graph is a straight line with a slope equal to the l_1-norm of the impulse response. (ii) Dead-beat anti-windup. Note that the response to data errors affecting the control signal (η_u) leads to larger control errors in this case. This is because the data error is interpreted as a saturation by the anti-windup, which is fast enough to react immediately. Important to note is also that the control error magnitude does not increase with the data error magnitude for large data errors. (iii) Dead-beat anti-windup only on the integrator state (this is common in e.g. PI-controllers). This is normally sufficient for handling actuator limitations, but in the case of data-errors the performance is inferior compared with full-state anti-windup. (iv) Dynamic anti-windup with a bandwidth of $2\omega_{cl}$. Here the response to data errors entering the control signal is better, since the anti-windup is too slow to react immediately on the error impulse. The response to data errors on the state is, however, worse than for the dead-beat design, as is clear from the figure. The recovery time is also significantly slower compared to the dead-beat design, which may be seen from time-domain response plots.

7. Discussion

In general, it can be noted that the integrator state z_1 is most sensitive to data errors, which seems very intuitive since it depends on feedback of the control error to decay, while the controller state z_2 is stable with a short time constant and decays by itself when the loop is broken by saturation. Another general observation is that the control-signal bound is computed from bounds on the command signal and, in consequence, is related to the expected magnitude of control errors during normal operation. Hence, the artificial limits and the anti-windup scheme will capture data errors resulting in control errors larger than those expected during normal operation. The proposed method may also be combined with the dynamic bounds of [Stroph and Clarke, 1999], to increase the coverage for data errors.

Since the controller state z is bounded by (10a), it may seem natural to introduce explicit limits also for this variable. However, simulations with saturations on the controller state indicate that the performance improvement is minor. The increased complexity resulting from additional saturations in the loop also makes the system difficult to analyze, even if it seems to perform well in presence of data errors.

In presence of stochastic signals, such as measurement noise, there is a probability of false error detections. This may be handled by computing the resulting variance of the control signal. The deterministic control-signal bound (10b) can then be adjusted with some measure depending

on the variance. The size of the adjustment will determine the probability of false detections.

Sporadic false detections will affect control performance, as the anti-windup then intervenes. By using slower anti-windup dynamics the noise sensitivity is decreased.

If a rate bound on the command signal is known $|\Delta u_c(k)| = |u_c(k) - u_c(k-1)| \leq M_\Delta$, then state and control-signal rate bounds may be computed in analogy with (10). An artificial rate bound on the control signal may then also be introduced in the controller, and used together with the anti-windup scheme. Note that the noise sensitivity will be worse than for the case with magnitude bounds, since the rate of the control signal will depend on the noise variance.

8. Summary

Data errors resulting from transient faults in the computer hardware may be modeled as impulse errors entering the control algorithm internally. The effect on closed-loop performance therefore depends on the controller realization. Previous methods to optimize controller realizations with respect to round-off errors in finite-precision numerics are applicable also in the context of transient data errors. Given a controller realization, robustness to data errors may be achieved by introducing artificial signal limitations based on l_∞-bounds, in combination with an anti-windup scheme. As one may expect, the integrator state of the controller is most sensitive to data errors.

References

Askerdal, Ö. (2000): *Design and Evalutation Techniques for Detection and Coverage Estimation of Low-Level Errors*. Lic. Thesis, Chalmers University of Technology, Sweden.

Askerdal, Ö., M. Gäfvert, M. Hiller, and N. Suri (2002): "A control theory approach for analyzing the effects of data errors in safety-critical control systems." In *Proceedings of the Pacific Rim International Symposium on Dependable Computing*.

Åström, K. J. and B. Wittenmark (1997): *Computer-Controlled Systems*, 3rd edition. Prentice Hall.

Constantinescu, C. (2002): "Impact of deep submicron technology on dependability of VLSI circuits." In *Proceedings of the IEEE International Conference on Dependable Systems end Networks*.

Cunha, J. C., R. Maia, M. Z. Rela, and J. G. Silva (2001): "A study of failure models in feedback control systems." In *Proceedings of the IEEE International Conference on Dependable Systems and Networks*, pp. 314–323.

Edwards, C. and I. Postlethwaite (1998): "Anti-windup and bumpless-transfer schemes." *Automatica*, **34:2**, pp. 199–210.

Hanselmann, H. (1987): "Implementation of digital controllers — A survey." *Automatica*, **23:1**, pp. 7–32.

Hazucha, P. (2000): *Background Radiation and Soft Errors in CMOS Circuits*. PhD thesis, Linköping University, Sweden.

Kim, H. and K. G. Shin (1994): "On the maximum feedback delay in a linear/nonlinear control system with input disturbances caused by controller-computer failures." *IEEE Transactions on Control Systems Technology*, **2:2**, pp. 110–122.

Kim, H., A. L. White, and K. G. Shin (2000): "Effects of electromagnetical interference on controller-computer upsets and system stability." *IEEE Transactions on Control Systems Technology*, **8:2**, pp. 351–357.

Moroney, P., A. S. Willsky, and P. Houpt (1983): "Roundoff noise and scaling in the digital implementaion of control compensators." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-31:6**, pp. 1464–1477.

Pradhan, D. (1996): *Fault-Tolerant Computer Systems Design*. Prentice Hall.

Rotea, M. A. and D. Williamson (1995): "Optimal realizations of finite wordlength digital filters and controllers." *IEEE Transactions on Circuits and Systems — I*, **42:2**, pp. 61–72.

Shivakumar, P., M. Kistler, S. W. Keckler, D. Burger, and L. Alvisi (2002): "Modeling the effect of technology trends on the soft error rate of combinatorial logic." In *Proceedings of the IEEE International Conference on Dependable Systems end Networks*.

Slegel, T. J., R. M. Averill III, M. A. Check, B. C. Giamei, B. W. Krumm, C. A. Krygowski, W. H. Li, J. S. Liptay, J. D. MacDougall, T. J. McPherson, J. A. Navarro, E. M. Schwarz, K. Shum, and C. F. Webb (1999): "IBM's S/390 G5 microprocessor design." *IEEE Micro*, **19:2**, pp. 12–23.

Stroph, R. and T. Clarke (1998): "Static acceptance test for complex controllers." In *Proceedings of the UKACC International Conference on CONTROL*, pp. 392–397. IEE.

Stroph, R. and T. Clarke (1999): "Dynamic fault detection approaches." In *Proceedings of the American Control Conference*, pp. 627–631.

Vinter, J. (2001): "Reducing critical failures for control algorithms using executable assertions and best effort recovery." In *Proceedings of the IEEE International Conference on Dependable Systems and Networks*, pp. 347–356.

Wittenmark, B., J. Nilsson, and M. Törngren (1995): "Timing problems in real-time control systems: Problem formulation." In *Proceedings of the American Control Conference*, pp. 2000–2004.