

Linköping Studies in Science and Technology
Dissertations, No 1589

Model Based Vehicle Level Diagnosis for Hybrid Electric Vehicles

Christofer Sundström



Linköping University
INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering
Linköping 2014

Linköping Studies in Science and Technology
Dissertations, No 1589

Christofer Sundström
`christofer.sundstrom@liu.se`
`www.vehicular.isy.liu.se`
Division of Vehicular Systems
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

Copyright © 2014 Christofer Sundström, unless otherwise noted.
All rights reserved.

Sundström, Christofer
Model Based Vehicle Level Diagnosis for Hybrid Electric Vehicles
ISBN 978-91-7519-356-4
ISSN 0345-7524

Hybrid powertrain illustration on the cover based on illustration by Lars Eriksson.

Typeset with L^AT_EX 2_ε
Printed by LiU-Tryck, Linköping, Sweden 2014

To Tilda, Elin and Siri

ABSTRACT

When hybridizing a vehicle, new components are added that need to be monitored due to safety and legislative demands. Diagnostic aspects due to powertrain hybridization are investigated, such as that there are more mode switches in the hybrid powertrain compared to a conventional powertrain, and that there is a freedom in choosing operating points of the components in the powertrain via the overall energy management and still fulfill the driver torque request. A model of a long haulage truck is developed, and a contribution is a new electric machine model. The machine model is of low complexity, and treats the machine constants in a different way compared to a standard model. It is shown that this model describes the power losses significantly better when adopted to real data, and that this modeling improvement leads to better signal separation between the non-faulty and faulty cases compared to the standard model.

To investigate the influence of the energy management design and sensor configuration on the diagnostic performance, two vehicle level diagnosis systems based on different sensor configurations are designed and implemented. It is found that there is a connection between the operating modes of the vehicle and the diagnostic performance, and that this interplay is of special relevance in the system based on few sensors.

In consistency based diagnosis it is investigated if there exists a solution to a set of equations with analytical redundancy, i.e. there are more equations than unknown variables. The selection of sets of equations to be included in the diagnosis system and in what order to compute the unknown variables in the used equations affect the diagnostic performance. A systematic method that finds properties and constructs residual generator candidates based on a model has been developed. Methods are also devised for utilization of the residual generators, such as initialization of dynamic residual generators, and for consideration of the fault excitation in the residuals using the internal form of the residual generators. For demonstration, the model of the hybridized truck is used in a simulation study, and it is shown that the methods significantly increase the diagnostic performance.

The models used in a diagnosis system need to be accurate for fault detection. Map based models describe the fault free behavior accurately, but fault isolability is often difficult to achieve using this kind of model. To achieve also good fault isolability performance without extensive modeling, a new diagnostic approach is presented. A map based model describes the nominal behavior, and another model, that is less accurate but in which the faults are explicitly included, is used to model how the faults affect the output signals. The approach is exemplified by designing a diagnosis system monitoring the power electronics and the electric machine in a hybrid vehicle, and simulations show that the approach works well.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Ett *diagnosystem* övervakar ett system för att fastställa om det är helt eller trasigt. Ett första steg är att upptäcka ett eventuellt fel, men det är även önskvärt att kunna peka ut vilken del av systemet som är trasigt. Övervakning av ett fordonets drivlina är viktigt av flera anledningar, bland annat för att uppfylla lagkrav, säkerhetskrav, hög utnyttjandegrad, och effektiva reparationer. När ett fordon hybridiseras, i den här avhandlingen genom att förbränningsmotorn kombineras med en elmaskin för fordonets framdrivning, så ökar systemets komplexitet och ställer därmed stora krav på det diagnosystem som övervakar fordonet. Det är vanligt att det finns ett diagnosystem för varje komponent i fordonets drivlina, men här studeras vilka fördelar det finns med att designa ett diagnosystem som övervakar ett flertal komponenter i fordonet. En speciell egenskap hos ett hybridiserat fordon är att det finns en frihet att välja om det är elmaskinen eller förbränningsmotorn som ska användas för att driva fordonet framåt. Därför är det intressant att studera hur designen av den övergripande energistyrningen påverkar möjligheten att felövervaka fordonet.

I avhandlingen används *konsistensbaserad diagnos*, vilket innebär att en matematisk modell över fordonet jämförs med sensorsignaler för att fastställa fordonets felstatus. För att undersöka hur olika designval påverkar diagnosprestandan har en modell av en lastbil utvecklats och ett bidrag i avhandlingen är en ny elmaskinmodell. Det visas att den nya modellen beskriver maskinens förluster bättre än en standardmodell när dessa utvärderas på mätdata, samt att modelleringsförbättringen leder till en bättre signalseparation mellan det felfria fallet och när ett fel har uppstått i systemet. Flera olika diagnosystem har designats och implementerats i simuleringsmodellen. Simuleringar visar bland annat att det finns en koppling mellan fordonets arbetspunkter och diagnosprestandan, samt att den kopplingen är av större betydelse när få sensorer är tillgängliga.

Grunden i de utvecklade diagnosystemen är att konstruera *residualgeneratorer*, som här undersöker om lastbilmodellen överensstämmer med sensormätningar. Det går att skapa tusentals residualgeneratorer baserat på en modell av ett komplext fysikaliskt system. Dessa har olika känslighet för att upptäcka fel i systemet, och därför har en metod som undersöker residualernas egenskaper baserat på en systemmodell utvecklats. Residualsignalerna i ett diagnosystem efterbehandlas och metoder för detta har konstruerats. En metod har även utvecklats för att kombinera en noggrann modell för det felfria fallet med en annan modell för samma fysikaliska system, men som beskriver hur de olika felen påverkar systemet. Detta leder till att det är möjligt att upptäcka fel i det övervakade systemet, och samtidigt även specificera vilken komponent som är felaktig, utan detaljerad modellering. För att demonstrera dessa metoder har en simuleringsstudie med lastbilmodellen utförts där det visas att metoderna signifikant förbättrar diagnosprestandan.

ACKNOWLEDGMENTS

First of all I would like to express my gratitude to my supervisor Professor Lars Nielsen for letting me join his research group and for all his support during these years. My second supervisor Erik Frisk is acknowledged for the many discussions about diagnosis and good comments for improving paper manuscripts.

Daniel Eriksson and Emil Larsson are acknowledged for different discussions about diagnosis in general throughout these years. Carl Svård is acknowledged for letting me use his implementation of the residual generator selection algorithm that is modified in Paper C, and Mattias Krysanter and Per Öberg for electric machine modeling discussions. During this period I have had the pleasure to collaborate with Daniel Eriksson, Mattias Krysanter, Xavier Llamas Comellas, Tomas Nilsson, Peter Nyberg, and Martin Sivertsson in different projects. Lars Eriksson is also acknowledged for involving me in an undergraduate course about vehicle propulsion in which the discussions with the students have given me many new insights about vehicle hybridization.

The industrial involvement in the project is valuable and Tobias Axelsson, Mattias Nyberg, Tobias Pettersson, Marcus Stigsson, and Nils-Gunnar Vågstedt are acknowledged for this. The colleagues at vehicular systems are acknowledged for creating a nice and pleasant atmosphere to work in. The never-ending discussions with Oskar Leufvén has given me insights in all from forestry to turbochargers, but most importantly lots of fun.

Finally, I would like to thank my family for all your support and encouragement. I would also like to express my gratitude to you Therése and to Tilda, Elin, and Siri, for bringing all your love and happiness into my life.

Contents

1	Introduction	1
1.1	Diagnosis	2
1.2	Outline and Contributions	4
1.3	Publications	6
	References	8
	Publications	11
A	Overall Monitoring and Diagnosis for Hybrid Vehicle Power-trains	13
1	Introduction	16
2	Simulation Platform	16
2.1	Environment	17
2.2	Driver model	18
2.3	Vehicle model	18
2.4	Controller and energy management	23
2.5	Driving cycles and simulation results	23
2.6	Sensors	24
2.7	Faults	24
3	Sensor configurations and theoretical maximum fault isolability .	27
3.1	Sensor configuration 1	27
3.2	Sensor configuration 2	28
4	Diagnosis systems design	28
4.1	Diagnosis system 1	30
4.2	Diagnosis system 2	32
5	Results and discussion	35
5.1	Diagnosis system 1	36
5.2	Diagnosis System 2	38
6	Conclusions	39
	References	41

B	A New Electric Machine Model and its Relevance for Vehicle Level Diagnosis	43
1	Introduction	46
2	Electric machine model	46
	2.1 Standard model	47
	2.2 New model	48
	2.3 Parametrization of the models	49
3	Vehicle Model	52
	3.1 Powertrain model	53
	3.2 Sensors	55
4	Evaluation on diagnosis system	56
	4.1 Induced fault	56
	4.2 Residuals	56
	4.3 Simulation results	57
5	Conclusions	60
	References	62
A	Notation	64
C	Selecting and Utilizing Sequential Residual Generators in FDI Applied to Hybrid Vehicles	67
1	Introduction	70
2	Residual generator construction	70
	2.1 Sequential Residual Generation by Structural Analysis . .	71
	2.2 An Algorithm	73
	2.3 Dynamic Models	73
	2.4 Modification to the Algorithm to Handle Dynamic Consistency Relations	74
3	Vehicle Model	75
	3.1 Powertrain model	75
	3.2 Sensors	78
	3.3 Induced faults	78
4	Selection of Consistency Relation	79
	4.1 Avoiding algebraic loops by consistency relation selection	79
	4.2 Properties of the sequential residual generators candidates	82
	4.3 Summary and discussion	83
5	Methods for Utilization of Residual Generators and Test Quantities	84
	5.1 Avoid differentiating in the consistency relation using a state transformation	84
	5.2 Initialization of states	86
	5.3 Consider fault excitation when computing test quantities	86
6	Illustrative Designs and Simulation Study	87
	6.1 Properties of diagnosis systems used in simulation study .	88
	6.2 Model used in the diagnosis system	90
	6.3 Initialization of states when restarting residual generators	90
	6.4 Two approaches for when to update dynamic test quantities	91

6.5	Simulations on driving cycle	94
6.6	Summing up	96
7	Conclusions	97
	References	99
A	Powertrain Model	102
B	Residual generators	103
B.1	Same tests in ICDS and MCDS	103
B.2	ICDS	103
B.3	MCDS	104
D	Diagnostic Method Combining Map and Fault Models Applied on a Hybrid Electric Vehicle	105
1	Introduction	108
1.1	Contributions and outline	108
2	Models of the electric machine	109
2.1	Map based model	109
2.2	Analytical model	110
3	Combining the map and analytical models for fault modeling	111
3.1	Finding an expression for ΔT_{em}	113
3.2	Finding an expression for $\Delta P_{em,l}$	114
4	Isolability gain by combining models	114
4.1	Theoretical fault isolability using map based model	114
4.2	Theoretical fault isolability using a combined model	115
5	Design of a diagnosis system	115
5.1	State space formulation of the model	116
5.2	Fault estimation	117
5.3	Design of residual generators and test quantities	121
5.4	Summing up	127
6	Conclusions	128
	References	129

Chapter 1

Introduction

There are possibilities to increase the efficiency of automotive powertrains using hybrid technology. The largest relative fuel saving can be obtained in city buses and garbage trucks with many start and stops, but also a small relative saving in the fuel consumption for long haulage trucks results in a large amount of saved fuel (Bradley, 2000). When hybridizing a vehicle, new components are added compared to a conventional vehicle, e.g. electric machines, battery pack, and power electronics (Husain, 2003; Guzzella and Sciarretta, 2013), and these components need to be accurately monitored (Diallo et al., 2013).

One reason for monitoring the powertrain is safety. Faults in the electrical components could be fatal due to the high voltage in the system. Another issue is to avoid that unintended torque is applied to the vehicle. Such faults are included in the functional safety of the vehicle, and there is an increased consideration to this field by the automotive industry. There is a global standard, ISO 26262, that provides processes and methods for the design, development, and manufacturing of vehicles, with the goal to determine the Automotive Safety Integrity Level (ASIL) in a systematic way (ISO, 2011). The ISO 26262 standard is not mandatory for heavy trucks, but this is likely to be changed by 2016 (Dardar et al., 2012).

In addition to safety, fault detection and isolation is important to decrease the vehicle ownership cost and to maximize the up-time of the vehicle. Fault detection can e.g. protect other components from breaking down if a fault occurs in a powertrain component. It is especially important to protect the battery that is expensive and may degrade fast (Chen et al., 2013), e.g. if there are large power flows in the battery. High power in the electrical components could be caused by a fault in the power electronics or the electric machine. Further, more efficient repair is possible if a fault is isolated, i.e. it is stated what component that is broken and to some extent also in what way. This leads to that the up-time of the vehicle increases, not only due to a minimization of the time

spent at the workshop, but also due to the possibility to implement fault tolerant control strategies in the vehicle.

Monitoring the powertrain of a hybrid electric vehicle (HEV) leads to new challenges. One example of this is that there are many different operating modes in an HEV. These operating modes offer possibilities to increase the performance of the diagnosis system, since there are at least two energy converters in the powertrain and thereby there is a freedom in choosing operating points of the components via the energy management.

1.1 DIAGNOSIS

A diagnosis scheme detects faults in a physical system using measurements, and there are several approaches to be used. In the process industry often data driven methods are used (Qin, 2012), while model based approaches e.g. are used in the automotive industry. Examples of model based approaches used in the control community are parity equation (Chow and Willsky, 1984), variable elimination (Staroswiecki and Comtet-Varga, 2001), parameter estimation (Isermann, 2006), state-observer (Frank, 1994), and residual generator (Blanke et al., 2006) techniques. From the AI field a common approach is consistency based diagnosis (de Kleer et al., 1992), that can be based on a general diagnostic engine (de Kleer and Williams, 1987; Struss and Dressler, 1989). An overview of the theories used in the control and AI communities is found in Travé-Massuyès (2014).

A diagnosis system using a model based approach uses a model of the monitored system, including a set of model equations E , describing the connection between the control and sensor signals for the nominal case. A residual generator included in the diagnosis system is designed based on a subset of the model equations, $\bar{E} \subseteq E$, with analytical redundancy, generically meaning there are more equations than unknown variables. One basic principle when constructing a residual generator based on \bar{E} , is that a subset $E' \subseteq \bar{E}$ of the equations is used to compute the unknown variables in \bar{E} , and the other equations, i.e. $\bar{E} \setminus E'$, are used to investigate the consistency between the model \bar{E} and the observations. Often the residual generators are based on a set of equations with analytical redundancy one, i.e. there is one more equation in \bar{E} than there are unknown variables in \bar{E} . The equations used to investigate the consistency between the model and the observations is called consistency relation or analytical redundancy relation (ARR) (Cassar and Staroswiecki, 1997; Staroswiecki and Comtet-Varga, 2001).

The computation of the unknown variables in a residual generator can be done by finding algebraic expressions for the variables or using numerical techniques, see e.g. (Brenan et al., 1996). One disadvantage using numerical solvers in nonlinear systems is that it is generally more computationally demanding compared to using algebraic expressions. Here, the designed diagnosis systems are supposed to be able to be implemented in a truck with limited computational

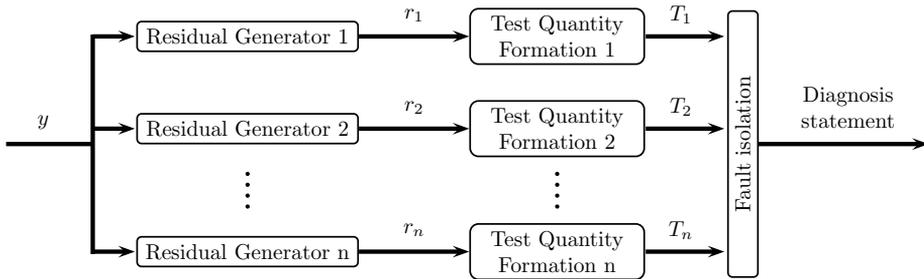


Figure 1.1: A typical diagnosis system includes several residuals that are post processed to form test quantities, and a fault isolation scheme to pinpoint which fault that has occurred.

power, and therefore algebraic expressions are found for the variables in the residual generators in Papers A-C, while the faults are estimated in state-observers in Paper D.

A diagnosis system often consists of several residual generators that are sensitive to different faults (Blanke et al., 2006). To reduce the noise level in the residual signals, these are post processed to form test quantities, as can be seen in Figure 1.1. The diagnosis statement is computed in a fault isolation scheme, where information about what test quantities that have reacted and what faults each test quantity is expected to react to is used.

STRUCTURAL ANALYSIS

When designing a diagnosis system the well known method called structural analysis can be used (Dustegör et al., 2006; Blanke et al., 2006; Staroswiecki and Declerck, 1989). The method is based on that all variables that are used in every model equation are listed, but no information about how the variables are included (e.g. linear, exponential, differentiated) is used. Using the structural analysis method it is possible to determine what detectability and isolability of the faults that are possible to ideally achieve given a model and a set of sensors (Krysander and Frisk, 2008).

The information about which variables that are included in each equation is included in a bipartite graph. Based on this graph a Dulmage-Mendelsohn decomposition (Dulmage and Mendelsohn, 1958) gives information about what part of the model that has analytical redundancy, and thereby can be monitored. There are several efficient tools available to find subsets of the model with analytical redundancy, and some of these are discussed and compared in Armengol et al. (2009).

Sets of model equations with analytical redundancy are of special interest when designing diagnosis systems, since they are used to construct residual generators, and are denoted ARRs, possible conflicts (Pulido and Gonzalez,

2004), and minimal structurally overdetermined (MSO) sets (Krysander et al., 2008) by different authors. A set of equations, \mathcal{M} , is structurally overdetermined if there are more equations than unknowns in \mathcal{M} . The set \mathcal{M} is an MSO set if there is no subset of \mathcal{M} that is structurally overdetermined. The structural method used when designing the diagnosis systems in Papers A-C are described in Krysander et al. (2008); Krysander and Frisk (2008).

VEHICLE LEVEL DIAGNOSIS

The manufacturers of the different components in a vehicle powertrain often also deliver a diagnosis system monitoring each component. When the components are connected in a hybrid powertrain it is however possible to design a diagnosis system monitoring the entire powertrain. This type of overall diagnosis is here called vehicle level diagnosis, and is the main emphasis of this thesis. There are several possible benefits of using such a diagnosis system. One benefit is that the performance of the diagnosis may increase, and another benefit is that it may be possible to monitor the components by using fewer sensors, compared to using separate diagnosis systems for each component in the powertrain.

1.2 OUTLINE AND CONTRIBUTIONS

The aim of this work is to investigate aspects influencing diagnosis on vehicle level regarding performance, design complexity, and computational complexity. Examples of such aspects are how the design of a diagnosis system affect performance, but also the importance of using accurate models for the purpose of diagnosis. A third example of an aspect is how the sensor configuration affects the diagnosis system. The aspects mentioned above are generic when designing diagnosis systems, but an aspect that is important to understand when monitoring HEVs is how the design of the energy management in combination with the driving mission either can hide or attenuate a fault. This aspect is of higher relevance in HEVs compared to conventional vehicles, since there are more mode shifts in the hybrid system, and there is a freedom in selecting operating modes of the powertrain components via the overall energy management. The understanding of such issues is crucial when constructing a diagnosis system on vehicle level for hybrid trucks.

Diagnostic aspects are investigated in the papers included in the thesis, and a summary of the contributions in each paper is presented below.

PAPER A

A simulation platform is used to evaluate the designed diagnosis systems in Papers A-D. The platform includes a vehicle powertrain model, possibility to induce faults in the powertrain, and a diagnosis system. Most of the powertrain component models are obtained from the existing MATLAB/SIMULINK model libraries CAPSIM (Fredriksson et al., 2006) and QSS (Guzzella and Amstutz,

1999), but are modified to represent a parallel hybrid truck as well as to include the possibility to add sensor noise and induce faults in the system. The simulation platform, with emphasis on the powertrain model, is described in Paper A and the model equations are given in the appendix of Paper C. In addition to the model description, in Paper A also two diagnosis systems are designed and implemented in the simulation platform. The two diagnosis systems are based on two different sensor configurations to investigate different aspects affecting the diagnosis of hybrid electric vehicles, such as how the choice of the sensor configuration affects the model based diagnosis system, but also the connection between the diagnostic performance and the operating modes of the vehicle. It is found that all faults are detected in both diagnosis systems, but full fault isolability is only achieved in the system based on more sensors. Further, there is a connection between the operating modes of the vehicle and the diagnostic performance, and this interplay is of special relevance in the system based on few sensors.

Paper A is a modified version of Sundström et al. (2010) extended with work presented in Sundström (2011).

PAPER B

When comparing the electric machine model used in Paper A, that also is described in Guzzella and Sciarretta (2013), to measurement data of an electric rear axle drive, it is found that the model does not capture the characteristics of the power losses in the machine. Therefore a new model of the electric machine is presented in Paper B. The model has low complexity to be able to use the model for on-board applications, such as in a diagnosis system. The new model treats the machine constants in a different way compared to the model described in Guzzella and Sciarretta (2013), which results in a different expression for the power losses. It is shown that the new model describes the power losses significantly better when adopted to real data compared to the standard model. The significance of the modeling improvement is demonstrated using a task in vehicle diagnosis where it is shown that the separation between the non-faulty and faulty cases is better and the resulting diagnostic performance is improved.

PAPER C

There are many residual generator candidates of a physical system, and a few of these are to be selected to be used in the diagnosis system. In Paper C the residual generators are based on MSOs, and all but one equation is used to compute the unknown variables and one equation is used to investigate the consistency between the model and the observations. A systematic method, that is based on Svärd and Nyberg (2010), to investigate the properties of the residual generators is described in the paper. The properties may differ between different residual generators, even these based on the same set of model equations, and therefore this kind of analysis is important. It may e.g. be

possible to find residual generators without algebraic loops, that are unique, or that either include differentiation or integration of dynamic equations. The algorithm proposed in Svärd and Nyberg (2010) is in Paper C extended to also consider the consistency relation in the analysis, and it is shown that only a small fraction of all residual generator candidates fulfill fundamental requirements, and thereby proves the value of such systematic methods. In addition, methods are devised for utilization of the residual generators, such as initialization of dynamic residual generators. A proposed method, considering the fault excitation in the residuals using the internal form of the residuals, significantly increases the diagnosis performance. The hybrid electric vehicle model is used in a simulation study for demonstration, but the methods used are general in character and provides a basis when designing diagnosis systems for other complex systems.

Paper C relies partly on work presented in Sundström et al. (2011).

PAPER D

High model accuracy directly results in good fault detection performance in a model based diagnosis system, and can be achieved by the use of a map based model. However, one drawback using such a model in a diagnosis system is the difficulty to isolate faults from each other, since internal physical phenomena are not described by the model. In Paper D a new diagnostic approach is presented to achieve also good fault isolability performance without extensive modeling. The map based model describes the nominal behavior of the monitored system, and another model, that is a less accurate but in which the faults are explicitly included, is used to model how the faults affect the output signals. The benefit of this approach is that data for a faulty system is not required, and that the accuracy demands on the model used for fault modeling are lower than for designing a diagnosis system without using the map based model. The approach is exemplified by designing an observer based diagnosis system monitoring the power electronics and the electric machine used in a hybrid electric powertrain, and simulations show that the approach works well.

Paper D relies partly on work presented in Sundström et al. (2013).

1.3 PUBLICATIONS

The research work leading to this thesis is presented in the following papers published by the author.

JOURNAL PAPERS

- C. Sundström, E. Frisk, and L. Nielsen. Selecting and utilizing sequential residual generators in FDI applied to hybrid vehicles. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, 44(2):172–185, February 2014 (**Paper C**)

SUBMITTED

- C. Sundström, E. Frisk, and L. Nielsen. A new electric machine model and its relevance for vehicle level diagnosis. 2014a. Submitted to Journal (**Paper B**)
- C. Sundström, E. Frisk, and L. Nielsen. Diagnostic method combining map and fault models applied on a hybrid electric vehicle. 2014b. Submitted to Journal (**Paper D**)

CONFERENCE PAPERS

- C. Sundström, E. Frisk, and L. Nielsen. Fault monitoring of the electric machine in a hybrid vehicle. In *7th IFAC Symposium on Advances in Automotive Control*, pages 548–553, Tokyo, Japan, 2013
- C. Sundström, E. Frisk, and L. Nielsen. Residual generator selection for fault diagnosis of hybrid vehicle powertrains. In *IFAC World Congress*, Milano, Italy, 2011
- C. Sundström, E. Frisk, and L. Nielsen. Overall monitoring and diagnosis for hybrid vehicle powertrains. In *6th IFAC Symposium on Advances in Automotive Control*, pages 119–124, Munich, Germany, 2010 (**Basis for Paper A**)

REFERENCES

- ISO 26262. Road Vehicles - Functional Safety. International Standard, 2011.
- J. Armengol, A. Bregon, T. Escobet, E. R. Gelso, M. Krysander, M. Nyberg, X. Olive, B. Pulido, and L. Trave-Massuyes. Minimal structurally overdetermined sets for residual generation: A comparison of alternative approaches. In *Proceedings of IFAC Safeprocess'09*, Barcelona, Spain, 2009.
- M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer, 2nd edition, 2006.
- R. Bradley. Technology roadmap for the 21st century truck program. Technical Report 21CT-001, U.S. Department of Energy, Oak Ridge, Tennessee, December 2000.
- K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Siam, 1996.
- J. Cassar and M. Staroswiecki. A structural approach for the design of failure detection and identification systems. In *Proceedings of IFAC Control of Industrial Systems*, Belfort, France, 1997.
- Z. Chen, Y. Fu, and C. Mi. State of charge estimation of lithium-ion batteries in electric drive vehicles using extended kalman filtering. *Vehicular Technology, IEEE Transactions on*, 62(3):1020–1030, 2013.
- E. Chow and A. Willsky. Analytical redundancy and the design of robust failure detection systems. *Automatic Control, IEEE Transactions on*, 29(7):603 – 614, July 1984.
- R. Dardar, B. Gallina, A. Johnsen, K. Lundqvist, and M. Nyberg. Industrial experiences of building a safety case in compliance with ISO 26262. In *IEEE 23rd International Symposium on Software Reliability Engineering Workshops (ISSREW)*, pages 349–354, 2012.
- J. de Kleer and B. C. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32:97–130, April 1987.
- J. de Kleer, A. Mackworth, and R. Reiter. Characterizing diagnoses and systems. *Artificial Intelligence*, 56(2-3):197–222, 1992.
- D. Diallo, M. Benbouzid, and M. Masrur. Special section on condition monitoring and fault accommodation in electric and hybrid propulsion systems. *Vehicular Technology, IEEE Transactions on*, 62(3):962–964, 2013.
- A. Dulmage and N. Mendelsohn. Coverings of bipartite graphs. *Canadian J. of Mathematics*, 10:517–534, 1958.

- D. Dustegör, E. Frisk, V. Coquempot, M. Krysander, and M. Staroswiecki. Structural analysis of fault isolability in the DAMADICS benchmark. *Control Engineering Practice*, 14(6):597–608, 2006.
- P. M. Frank. Enhancement of robustness in observer-based fault detection. *International J. of Control*, 59(4):955–981, 1994.
- J. Fredriksson, J. Larsson, J. Sjöberg, and P. Krus. Evaluating hybrid electric and fuel cell vehicles using the capsim simulation environment. In *22nd International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exposition*, pages 1994–2004, Yokohama, Japan, 2006.
- L. Guzzella and A. Amstutz. CAE tools for quasi-static modeling and optimization of hybrid powertrains. *IEEE Trans. on Vehicular Technology*, 48(6):1762–1769, November 1999.
- L. Guzzella and A. Sciarretta. *Vehicle Propulsion System, Introduction to Modeling and Optimization*. Springer Verlag, Zürich, 3 edition, 2013.
- I. Husain. *Electric and Hybrid Vehicles*. CRC Press LLC, 2003.
- R. Isermann. *Fault Diagnosis Systems - An Introduction from fault Detection to Fault Tolerance*. Springer Verlag, 2006.
- M. Krysander and E. Frisk. Sensor placement for fault diagnosis. *IEEE Trans. on SMC – Part A*, 38(6):1398–1410, 2008.
- M. Krysander, J. Åslund, and M. Nyberg. An efficient algorithm for finding minimal over-constrained sub-systems for model-based diagnosis. *IEEE Trans. on SMC – Part A*, 38(1), 2008.
- B. Pulido and C. Gonzalez. Possible conflicts: a compilation technique for consistency-based diagnosis. *IEEE Trans. on SMC – Part B*, 34(5):2192–2206, October 2004.
- S. J. Qin. Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36(2):220–234, 2012. ISSN 1367-5788.
- M. Staroswiecki and G. Comtet-Varga. Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems. *Automatica*, 37(5):687–699, 2001.
- M. Staroswiecki and P. Declerck. Analytical redundancy in non-linear interconnected systems by means of structural analysis. In *Proceedings of IFAC AIPAC'89*, pages 51–55, Nancy, France, 1989.
- P. Struss and O. Dressler. "physical negation": integrating fault models into the general diagnostic engine. In *Proceedings of the 11th international joint conference on Artificial intelligence - Volume 2*, pages 1318–1323, San Francisco, CA, 1989.

- C. Sundström. Vehicle level diagnosis for hybrid powertrains. Technical report, 2011. Licentiate thesis LiU-TEK-LIC-2011:27, Thesis No. 1488.
- C. Sundström, E. Frisk, and L. Nielsen. Overall monitoring and diagnosis for hybrid vehicle powertrains. In *6th IFAC Symposium on Advances in Automotive Control*, pages 119–124, Munich, Germany, 2010.
- C. Sundström, E. Frisk, and L. Nielsen. Residual generator selection for fault diagnosis of hybrid vehicle powertrains. In *IFAC World Congress*, Milano, Italy, 2011.
- C. Sundström, E. Frisk, and L. Nielsen. Fault monitoring of the electric machine in a hybrid vehicle. In *7th IFAC Symposium on Advances in Automotive Control*, pages 548–553, Tokyo, Japan, 2013.
- C. Sundström, E. Frisk, and L. Nielsen. Selecting and utilizing sequential residual generators in FDI applied to hybrid vehicles. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, 44(2):172–185, February 2014.
- C. Sundström, E. Frisk, and L. Nielsen. A new electric machine model and its relevance for vehicle level diagnosis. 2014a. Submitted to Journal.
- C. Sundström, E. Frisk, and L. Nielsen. Diagnostic method combining map and fault models applied on a hybrid electric vehicle. 2014b. Submitted to Journal.
- C. Svärd and M. Nyberg. Residual generators for fault diagnosis using computation sequences with mixed causality applied to automotive systems. *IEEE Trans. on SMC – Part A*, 40(6):1310–1328, 2010.
- L. Travé-Massuyès. Bridging control and artificial intelligence theories for diagnosis: A survey. *Engineering Applications of Artificial Intelligence*, 27(0):1 – 16, 2014.

Publications

Overall Monitoring and Diagnosis for Hybrid
Vehicle Powertrains*

A

*Revised and expanded version of a paper entitled *Overall Monitoring and Diagnosis for Hybrid Vehicle Powertrains* presented at *6th IFAC Symposium on Advances in Automotive Control, Munich, Germany, 2010*.

Overall Monitoring and Diagnosis for Hybrid Vehicle Powertrains

Christofer Sundström, Erik Frisk, and Lars Nielsen

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, SE-581 83 Linköping, Sweden.*

ABSTRACT

Designing diagnosis systems for hybrid vehicles includes new features compared to conventional vehicles, e.g., additional mode switches in the system. Aspects affecting diagnosis of hybrid electric vehicles are to be studied, and a main topic is a study on how the choice of the sensor configuration affects the model based diagnosis system, but also the connection between the diagnostic performance and the operating modes of the vehicle. These aspects are investigated by designing and implementing two diagnosis systems on vehicle level that are based on two sensor configurations, one consisting as few sensors as possible that theoretically achieve full fault isolability, and one having more sensors. The diagnosis systems detect specific faults, here faults in the electrical components in a hybrid electric powertrain, but the presented methodology is generic. It is found that all faults are detected in both diagnosis systems, but there is a connection between the operating modes of the vehicle and the diagnostic performance, and this interplay is of special relevance in the system based on few sensors. This leads to that it is possible to reduce the number of sensors used in the vehicle, if the diagnostic performance is considered when the overall energy management is designed.

1 INTRODUCTION

When hybridizing a vehicle, new components are added compared to a conventional vehicle, e.g. electric machines, battery, and power electronics (Husain, 2003; Guzzella and Sciarretta, 2013). These components need to be monitored with the same accuracy as the components used in a conventional vehicle, and one reason for monitoring the system is safety. Other reasons are to minimize the cost of the vehicle ownership, and to maximize the up-time of the vehicle, that is especially important in commercial vehicles. Accurate diagnosis leads to more efficient repair at the workshop and thereby lower cost, but also to the possibility to protect components, especially the battery that is sensitive and costly (Reddy, 2011; Chen et al., 2013), from breaking down if a fault occurs.

Monitoring a hybrid electric vehicle (HEV) powertrain leads to new challenges compared to a conventional powertrain. There are e.g. many different operating modes in an HEV, and one example is that the electric components are either active or not. In an HEV there is a freedom in choosing operating points of the components via the overall energy management of the vehicle, which is not possible to do in a conventional vehicle.

The objective of this work is to study key topics for vehicle level monitoring and diagnosis of hybrid vehicles. A main topic is a study on how the choice of the sensor configuration affects the model based diagnosis system, but also the connection between the diagnostic performance and the operating modes of the vehicle, and the influence of using a model of a component that is not valid in all operating modes. The presentation starts with a thorough description of the used simulation environment in Section 2, and the vehicle model parameters are set to represent a long haulage truck. To evaluate the connection between sensor configuration and diagnosis performance, two different sensor configurations are assumed available, and these are given in Section 3. Based on these two sensor configurations, two diagnosis systems are designed in Section 4. The results from a simulation study are presented in Section 5, and finally the conclusions are given in Section 6.

2 SIMULATION PLATFORM

To investigate the interplay between vehicle, controller, and driver with emphasis on fault monitoring and diagnosis, a simulation platform in MATLAB/SIMULINK has been developed. The simulation platform includes descriptions of the truck, driver model, controller and energy management algorithms, and different diagnosis systems. The diagnosis framework used in the paper is consistency based diagnosis using precompiled tests, or residuals, see for example Blanke et al. (2006) or the references therein. For logical foundations of the approach, see for example de Kleer et al. (1992).

The structure of the platform is given in Figure 1. The vehicle model is based on models of the components with a fixed interface to be able to easily

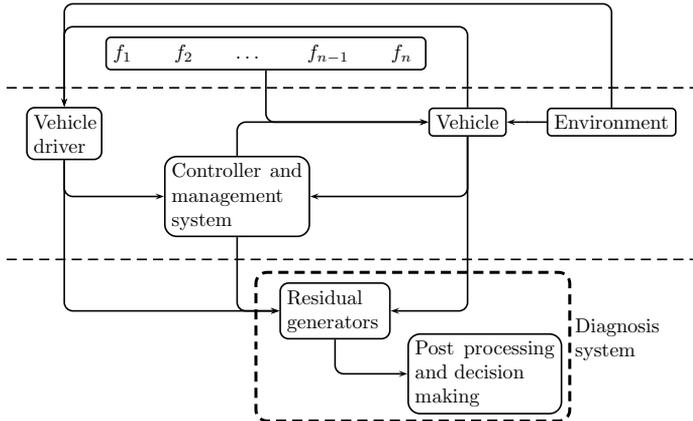


Figure 1: The structure of the implemented platform. Above the upper horizontal line the faults are modeled, between the lines the models of the vehicle, driver, and environment, and the lowest level contains the diagnosis system described in Sections 3 and 4.

change a component model without modifying the rest of the vehicle model. The diagnosis systems consist of two parts. First residuals, r , are computed, which are signals that ideally are zero in the fault free case and non-zero in a faulty case. To detect whether a residual significantly differs from zero, i.e. there is a fault in the system, the residual signals are post processed. Different residuals are sensitive to different faults, and the pattern of which residuals that are significantly non-zero and zero is used to state what part of the system that is faulty. In Figure 1, the lower dashed line has a special meaning in that the levels above are simulated as a time-continuous system, whereas the level below is simulated at a fix sample rate to represent vehicle on-board execution of the diagnosis system. In the platform, the models of the vehicle, environment and driver are based on the model libraries developed in the Center for Automotive Propulsion Simulation, CAPSIM, (Fredriksson et al., 2006) and QSS (Guzzella and Amstutz, 1999). Modifications to these models are carried out to model a truck instead of a passenger car, but also to include the possibility to induce faults in the models and to add sensor noise. The components in Figure 1 that are above the lower dashed line are described in this section, and the diagnosis systems in Sections 3 and 4.

2.1 ENVIRONMENT

The environment contains information about the driving cycle, i.e. the speed profile and the road gradient. The driving cycles used are presented in Section 2.5, where also simulations of the vehicle are carried out.

2.2 DRIVER MODEL

The model of the driver is a PI-controller using the information from the actual speed and the reference speed from the driving cycle, to set the position of the accelerator and brake pedals. The gear shifting strategy depends on the vehicle speed (see Sundström (2011) for details), and when to engage or disengage the clutch is also handled by the driver model.

2.3 VEHICLE MODEL

The vehicle modeled is a long haulage truck with a total weight of 40.000 kg, and the configuration of the powertrain is a parallel hybrid according to Figure 2.

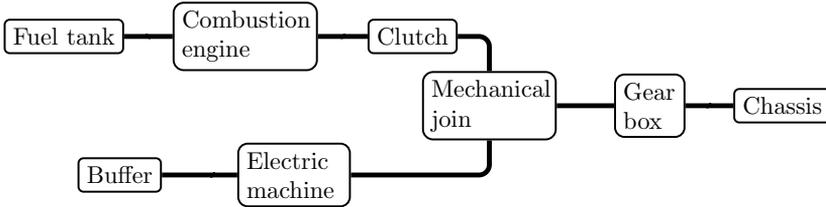


Figure 2: The modeled truck is a parallel hybrid with the connection of the electrical and conventional parts of the powertrain between the clutch and the gearbox.

Since the objective with this work is to study the interaction between the components in the vehicle, it is preferable to use basic component models. It is however easy to add more advanced models of the components. The used models of the components in the powertrain is briefly described below, and a more detailed description is given in Sundström (2011).

BATTERY

The battery is modeled as a voltage source, U_{oc} , and an inner resistance, R_b , connected in series (Reddy, 2011). The battery current, I_b , is expressed as

$$U_b = U_{oc}(SoC) - R_b I_b \quad (1)$$

and U_{oc} , varies with the state of charge, SoC , that is defined by

$$SoC = SoC_0 - \frac{1}{Q_b} \int I_b dt, \quad SoC \in [0, 1] \quad (2)$$

where SoC_0 is the initial state of charge and Q_b is the battery capacity.

The Coulombic efficiency is assumed to be negligible since this efficiency is close to one in lithium-ion batteries (Valøen and Shoesmith, 2007). The storage capacity is 9 kWh and U_{oc} is assumed to be constant, 256 V, when $SoC \in [0.2, 0.8]$.

ELECTRIC MACHINE

The electric machine is able to convert electric power to mechanical power and vice versa. A voltage, U_{em} , is applied on the component, resulting in a torque on the outgoing shaft. The torque, T_{em} , is modeled to be proportional to the armature current, I_{em} , with the torque constant k_a

$$T_{em} = I_{em}k_a \quad (3a)$$

The speed constant, k_i , is used to model the electromotive force, and R_{em} is the armature resistance of the machine

$$I_{em} = \frac{U_{em} - k_i\omega_{em}}{R_{em}} \quad (3b)$$

The model is parametrized as a 33 kW separately excited DC machine with constant magnetic flux, and the parameter values of R_{em} , k_a , and k_i are set to 0.044 Ω , 0.50 Nm/A , and 0.51 Vs/rad , respectively. In an ideal machine, k_a and k_i are equal, and are defined by $K\Phi$, where K is a machine constant that depends on the design parameters of the machine, and Φ is the magnetic flux produced by the stator (Guzzella and Sciarretta, 2013).

LOCAL CONTROLLER OF THE ELECTRIC MACHINE

The controller of the machine sets a requested voltage $U_{em,ctrl}$ to be applied on the machine by the power electronics. This is done by using a feed forward controller based on the model of the machine presented above. The voltage required to achieve the requested torque, $T_{em,req}$, set in the energy management is computed by

$$U_{em,ctrl} = \frac{R_{em}}{k_a} \left(T_{em,req} + \frac{k_i k_a}{R_{em}} \omega_{em} \right) \quad (4)$$

that is based on (3a) and (3b), and T_{em} is replaced by $T_{em,req}$.

POWER ELECTRONICS

The power electronics is modeled to deliver the requested voltage from the local controller of the electric machine, i.e.

$$U_{em} = U_{em,ctrl} \quad (5)$$

The component is assumed to be ideal

$$P_b = P_{em} \iff I_b U_b = I_{em} U_{em} \quad (6)$$

where P_b and P_{em} are the electrical powers from the battery and the machine.

FUEL TANK

In the model of the fuel tank the mass of the fuel in the tank, m_f , is computed by integrating the fuel mass-flow, \dot{m}_f , to the engine. The integrator is initialized with the mass of the fuel at the beginning of the driving cycle, $m_{f,0}$

$$m_f = \int -\max\{0, \dot{m}_f\} dt + m_{f,0} \quad (7)$$

The weight reduction of the vehicle when fuel is consumed is also computed and is used to compute the weight of the vehicle used in the model of the chassis

$$m_{f,r} = \int \max\{0, \dot{m}_f\} dt \quad (8)$$

ENGINE

The engine model is a mean value model that computes the delivered torque, T_e , by using the mean brake effective pressure, p_{me} . The mean brake effective pressure is defined as

$$p_{me} = \frac{4\pi T_e}{V_d} \quad (9)$$

where V_d is the displacement of the engine. The pressure p_{me} is calculated using Willans approximation (Guzzella and Sciarretta, 2013; Rizzoni et al., 1999). The indicated engine efficiency, $\eta_{e,i}$, i.e the efficiency of the transformation from chemical energy to pressure inside the cylinders, and the pumping and friction losses, p_{me0} , are considered

$$p_{me} = \eta_{e,i} p_{m\phi} - p_{me0} \quad (10)$$

The pressure p_{me0} can be divided into the the pumping losses, $p_{me0,g}$ and the friction losses, $p_{me0,f}$

$$p_{me,0} = p_{me0,f} + p_{me0,g} \quad (11)$$

where the pumping losses are assumed to be constant. The friction losses, $p_{me0,f}$, are modeled using a friction model given in Guzzella and Onder (2004), that is a simplified model of Inhelder (1996). In the expression, $k_{\{1,2,3,4\}}$ are constants, B and S the bore and stroke, and Π_{bl} the boost layout of the engine that affects the dimensioning of e.g. bearings

$$p_{me0,f} = k_1(k_2 + k_3 S^2 \omega_e^2) \Pi_{bl} \sqrt{\frac{k_4}{B}} \quad (12)$$

The parameters are based on Volvo's D16 that produces 700 hp. General parameters in the Willans approximation such as the indicated efficiency are the same that are used for a diesel engine in QSS (Guzzella and Amstutz, 1999). Some of the parameters used are presented in Table 1.

Table 1: Some key parameters used in the internal combustion engine

Parameter	Value	Unit
Number of cylinders	6	[-]
Stroke	0.165	[m]
Bore	0.144	[m]
Indicated efficiency	0.50	[-]
Max torque (speed)	3150 (1250)	[Nm (rpm)]
Max power (speed)	515 (1700)	[kW (rpm)]
Mass	800	[kg]

CLUTCH

There is a model of the clutch to handle starts and gear shifts. The model is included in CAPSIM, and when the clutch is engaged or disengaged, the component is modeled as an ideal component. A flywheel is included in the model and the difference in angular speed of the flywheel and the outgoing shaft is computed. This difference in speed is used to find the outgoing torque from the component when the clutch is not fully engaged.

MECHANICAL JOINT

The mechanical joint in Figure 2 connects the shafts from the electric machine, the clutch, and the gearbox. The torque delivered from the component, which is the torque on the input shaft to the gearbox, is denoted T_{mj} . There is a gear ratio, u_{em} , that is applied between the shaft connected to the electric machine and the other shaft connected to the clutch and gearbox.

$$T_{mj} = T_e + T_{em}u_{em} \quad (13)$$

The inertia is calculated using

$$J_{mj} = J_e + J_{em}u_{em}^2 \quad (14)$$

where J_e and J_{em} are the inertia of the engine and electric machine used to compute the acceleration of the vehicle in Section 2.3.

GEAR BOX

A fix step manual gearbox is used in the powertrain. The used gear is an input signal to the gearbox and is set in the vehicle driver model. Based on this signal the gear ratio, u_{gb} , is achieved. The losses in the gearbox are modeled using an affine dependency between the input and output torques. The torque consumed at idle is denoted as $T_{gb,l}$, and the proportional coefficient is denoted as η_{gb} , and

how η_{gb} is included in the expression depends on the sign of the delivered torque, T_{gb} from the gearbox

$$T_{gb} = \begin{cases} u_{gb} (T_{mj} - T_{gb,l}) \eta_{gb} & T_{mj} - T_{gb,l} \geq 0 \\ u_{gb} (T_{mj} - T_{gb,l}) \frac{1}{\eta_{gb}} & T_{mj} - T_{gb,l} < 0 \end{cases} \quad (15)$$

The efficiency η_{gb} depends on the selected gear, and $T_{gb,l}$ depends on the in-going speed and the selected gear. The inertia from the input shaft is compensated for the gear ratio when the inertia of the vehicle is calculated

$$J_{tot} = J_{gb} + u_{gb}^2 J_{mj} \quad (16)$$

The gearbox model is parametrized to represent Volvo's Ishift with 12 gears. The gear ratios vary between 11.73 (1st gear) and 0.78 (12th gear), and the efficiency $\eta_{gb} = 0.975$.

CHASSIS

In the vehicle, that has a mass m_v , the output shaft from the gearbox is connected to the final gear and finally to the wheels. The vehicle mass is computed by reducing the initial mass of the vehicle, $m_{v,0}$ with the consumed fuel, $m_{f,r}$

$$m_v = m_{v,0} - m_{f,r} \quad (17)$$

The road slope is used to calculate the change in potential energy, but is not used in the expression for the rolling resistance. The rolling resistance is modeled using the coefficient C_r

$$F_r = m_v g C_r \quad (18)$$

and g is the gravity constant. To be able to handle low velocities and stand still, the torque due to the rolling resistance, T_r , is modeled to be proportional to the angular speed of the wheels, ω_w , at low speeds. If the vehicle is reversing, T_r changes sign in the model

$$T_r = \begin{cases} F_r r_w, & 1000\omega_w > F_r r_w \\ 1000\omega_w, & -F_r r_w \leq 1000\omega_w < F_r r_w \\ -F_r r_w, & 1000\omega_w \leq -F_r r_w \end{cases} \quad (19)$$

The expression for the air drag torque includes the air density, ρ , the air drag coefficient, C_d , the frontal area of the vehicle, A_f , the vehicle speed, v , and the wheel radius

$$T_d = \frac{1}{2} \rho C_d A_f \omega_w^2 r_w^3 \quad (20)$$

The torque due to the road slope, θ , and thereby change in potential energy of the vehicle is modeled as

$$T_g = m_v g r_w \sin \theta \quad (21)$$

Table 2: The parameters used in the model of the chassis.

Parameter	Value	Unit
$m_{v,0}$	40000	[kg]
Tire specification	315/80R22.5	[-]
C_r	0.007	[-]
C_d	0.8	[-]
A_f	10	[m ²]
u_f	3.21	[-]

The torque due to mechanical brakes is denoted as T_B , and the gear ratio in the final gear is denoted as u_f . The net torque acting on the wheels is computed by

$$T_{net} = T_{gb}u_f - T_d - T_B - T_r - T_g \quad (22)$$

The effective inertia and the mass of the vehicle are used to calculate the angular acceleration of the wheels

$$\dot{\omega}_w = \frac{T_{net}}{J_{tot}u_f^2 + m_v r_w^2} \quad (23)$$

Some of the chassis parameters used in the simulations in Section 5 are given in Table 2.

2.4 CONTROLLER AND ENERGY MANAGEMENT

There are several design approaches of the energy management, e.g. the global optimal solution using dynamic programming (Lin et al., 2003), model predictive control (Borhan et al., 2009), or finding equivalent-consumption minimization strategies (ECMS) (Sciarretta and Guzzella, 2007; Sivertsson et al., 2011). In this study a heuristic approach is used since it is less complex than the above mentioned methodologies, and the focus is here on the design of the diagnosis systems.

The basic idea in the developed energy management is to charge the battery via the electric machine instead of using the mechanical brakes and later use this energy as a complement to the combustion engine. To be able to store as much energy as possible during a retardation, it is preferred to have a low state of charge, SoC , in the battery before the retardation. This is achieved by primarily propel the vehicle via the electrical machine if SoC is larger than a predefined threshold, SoC_{low} . The details in the design of the energy management are given in Sundström (2011).

2.5 DRIVING CYCLES AND SIMULATION RESULTS

Simulations of the vehicle are carried out to verify the model and to evaluate the diagnosis systems designed in Section 4. Two driving cycles are used, FTP75 and

a velocity profile collected from real driving between Linköping and Jönköping. FTP75 is a driving cycle including many starts and stops (see Figure 3), while the collected data represents highway driving. As seen in Figure 4, the truck is driving at constant speed during most of the time at Linköping-Jönköping, but at a few times the vehicle decreases the velocity. At a few downhill slopes of the road the vehicle brakes to keep constant speed. When this occurs the battery is charged, which can be seen in the figure. The fuel consumption is 39 l/100km when driving from Linköping to Jönköping, which is a reasonable fuel consumption for a fully loaded long haulage truck (Bradley, 2000).

Diagnosis of the electrical parts of the powertrain is of high interest in this paper. With the designed energy management, these components are only active if there are some energy to recuperate, or there are energy stored in the batteries. The electrical components are frequently active when FTP75 is used, since this driving cycle includes many starts and stops. When diagnosis systems are evaluated using the simulation model, it may be preferable to use a driving cycle that frequently excites the components that are to be monitored, and FTP75 is mainly used for this purpose. The recorded data from Linköping to Jönköping is used to verify that the results are valid for a long haulage truck also in more standard highway driving.

2.6 SENSORS

The truck is assumed to be equipped with sensors measuring voltages, currents, rotational speeds, and torques. The sensors measure the signals at 80 Hz, and noise is added to the sensor signals. The noise is assumed to be additive independent and identically distributed (i.i.d.) Gaussian with a standard deviation σ_i , and is denoted $\tilde{\nu} \in \mathcal{N}(0, \sigma_i)$. The measurement signal y is given by the noise free signal y^* , added with $\tilde{\nu}$ as

$$y = y^* + \tilde{\nu} \quad (24)$$

2.7 FAULTS

To model that the battery, power electronics, or the electric machine may break down, two parameter values and two voltages in these models have the possibility to be modified. Note that these faults are only examples of how a fault in these components can be represented in the model. The following modifications of the signals are introduced to model the faults and the nominal signals are denoted by the superscript *nom*:

$$f_{em,ka} : k_a = (1 + f_{em,ka}) k_a^{nom} \quad (25a)$$

$$f_{em,R} : R_{em} = (1 + f_{em,R}) R_{em}^{nom} \quad (25b)$$

$$f_{pe} : U_{em,ctrl} = (1 + f_{pe}) U_{em,ctrl}^{nom} \quad (25c)$$

$$f_{b,sc} : U_b = (1 + f_{b,sc}) U_b^{nom} \quad (25d)$$

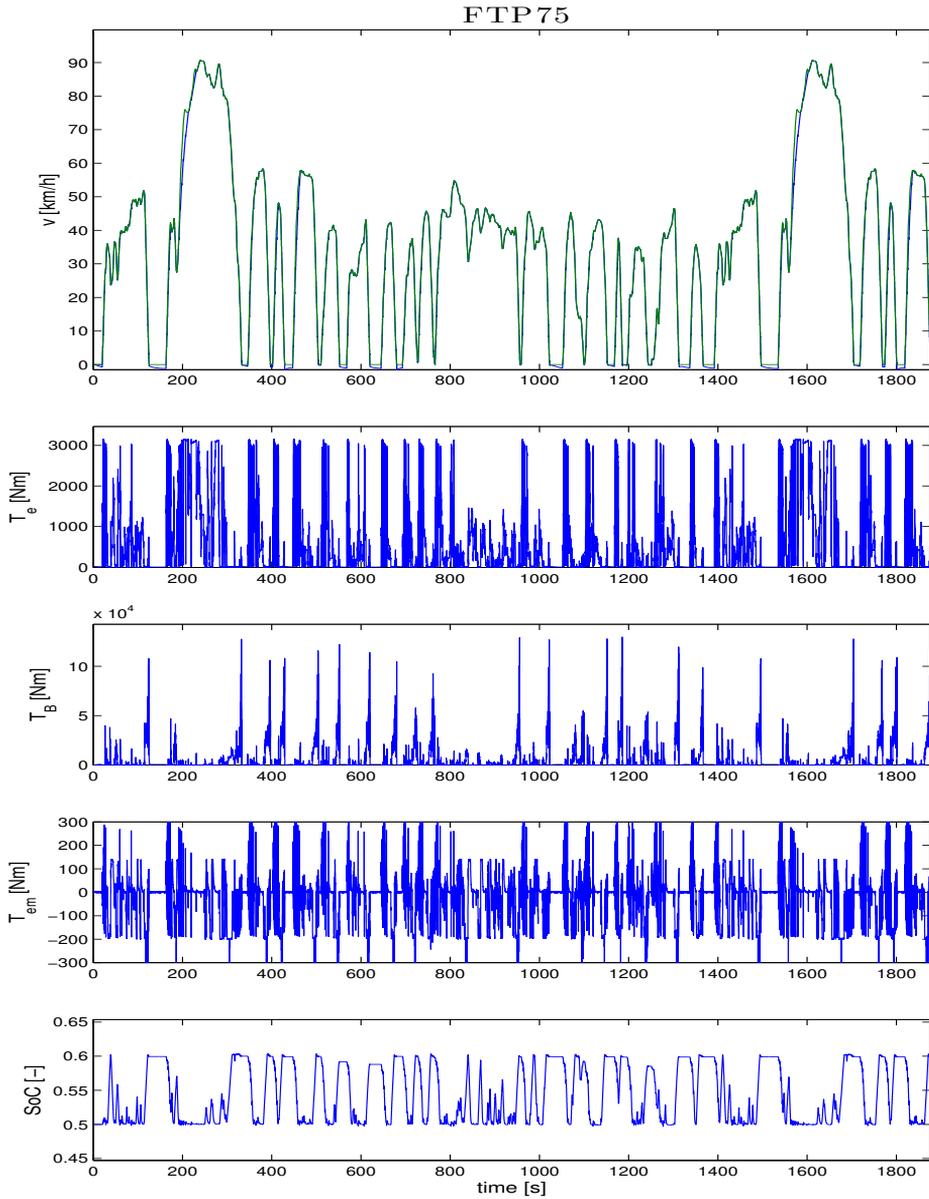


Figure 3: The reference velocity and the velocity of the vehicle when FTP75 is used, are given in the first plot. The engine, brake, and electric machine torques, as well as the SoC of the battery are also presented.

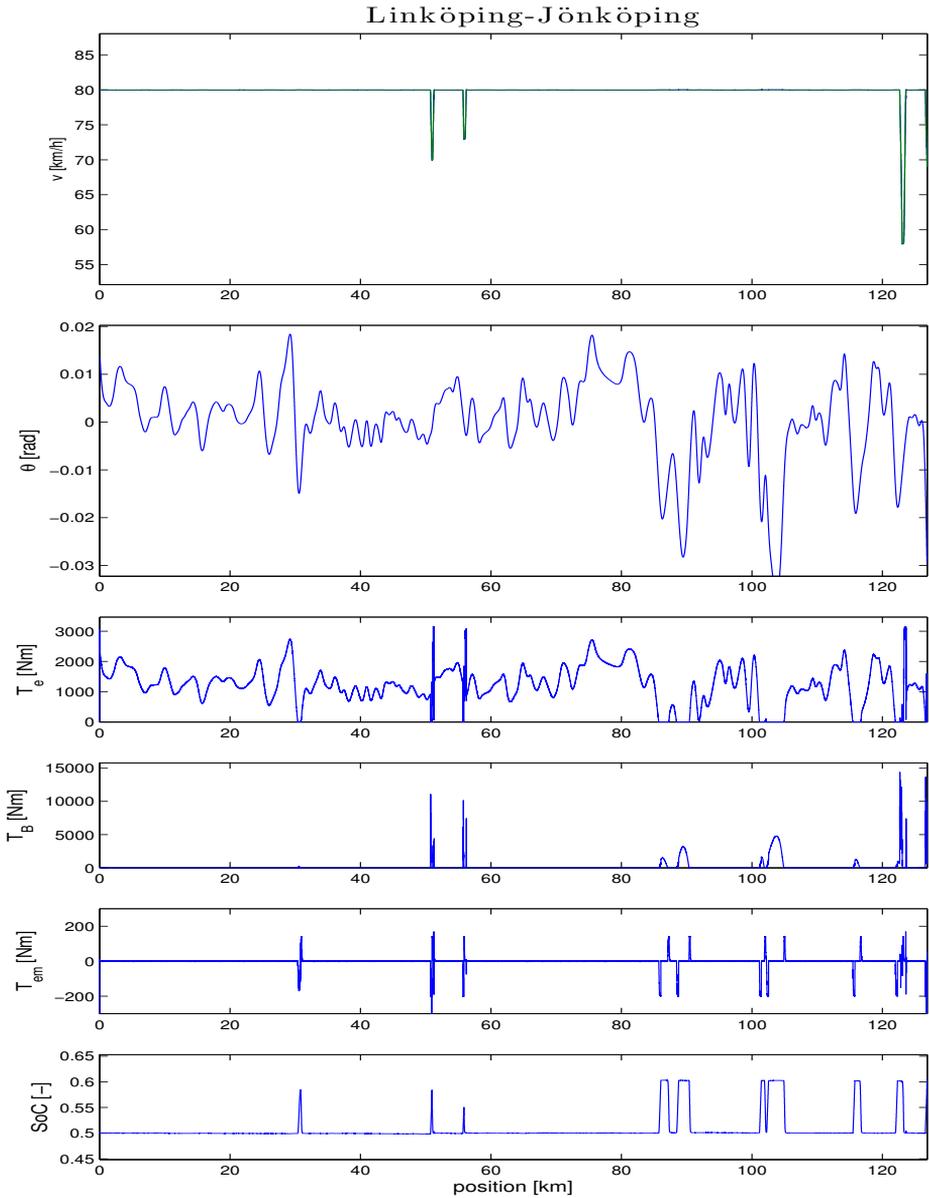


Figure 4: The velocity of the truck and road slope when driving from Linköping to Jönköping are presented in the upper plots. The engine, brake, and electric machine torques, as well as SoC are also shown. The electric machine is not used during long periods in this driving scenario.

Table 3: Values of faults induced in the model. The voltage U_{em} varies in the range 0 – 200 V, $U_b \approx 250$ V and $\omega_{gb} < 50$ rad/s.

Faults	Value
$f_{em,ka}$	-0.5
$f_{em,R}$	-0.5
f_{pe}	-0.5
$f_{b,sc}$	-0.5
$f_{b,U,sens}$	20 V
$f_{em,U,sens}$	20 V
$f_{\omega,gb,sens}$	20 rad/s

where $f_{b,sc}$ models that not all cells in the battery are used due to an internal short circuit, f_{pe} is a fault in the power electronics, and $f_{em,ka}$ and $f_{em,R}$ are two faults in the electric machine.

Three sensor faults are modeled as an offset fault, e.g. for the voltage sensor in the electric machine

$$U_{em,sens} = U_{em} + f_{em,U,sens} \quad (25e)$$

where $f_{em,U,sens}$ possibly is time variant. The other sensor faults are $f_{b,U,sens}$, and $f_{\omega,gb}$, that are faults in a battery voltage sensor and the speed sensor in the gearbox respectively.

When a fault is induced in the model, the value of the fault is given in Table 3.

3 SENSOR CONFIGURATIONS AND THEORETICAL MAXIMUM FAULT ISOLABILITY

To analyze how the choice of sensor configuration affects the performance and complexity of the diagnosis system, two systems using different sensor configurations are developed. One set of sensors is chosen to achieve a diagnosis system that is easy to design, and the other set is chosen to use as few sensors as possible. The sensor configurations will also be investigated to see how the sensor noise affects the diagnostic performance.

3.1 SENSOR CONFIGURATION 1

The faults to be detected and isolated in the first diagnosis system are the four faults described in (25a)-(25d), and the sensor fault described in (25e). This diagnosis system uses sensors that measure signals close to the components that are to be monitored, i.e. the battery, electric machine and power electronics:

- $\omega_{gb,sens}$ - gear box outgoing speed

- $I_{b,sens}$ - battery current
- $I_{em,sens}$ - electric machine current
- $T_{em,sens}$ - electric machine torque
- $U_{em,sens}$ - electric machine voltage

Note that in this system a torque sensor in the electric machine is used. Torque sensors are normally not used in series production vehicles, but in this case the torque sensor is included in the system to investigate its impact on the diagnostic performance.

Given a model and a set of sensors it is possible to determine what detectability and isolability of the faults that are theoretically possible to achieve. In Krysander and Frisk (2008) this is done by a structural analysis (Dustegör et al., 2006; Blanke et al., 2006) of the model. The method is based on that all variables that are used in every equation are listed. How the variables are included (e.g. linear, exponential, differentiated) is not considered in this analysis. The structural model using the above described sensor configuration is shown in Figure 5, where the last five equations represent the sensor equations and are modified if a different sensor configuration is used. Given the set of sensors described above it is possible to structurally achieve full fault isolability of the faults described in Section 2.7.

3.2 SENSOR CONFIGURATION 2

In the second diagnosis system the faults given in (25a)-(25d), in addition to faults in all used sensors, are to be detected and isolated. The number of sensors used to achieve this is minimized to be able to analyze the impact this choice has on the performance of the diagnosis system. To find this set of sensors, a sensor placement algorithm (Krysander and Frisk, 2008) is run using the structural model. Three sensors are required to isolate these faults, but there are several different sets with three sensors that structurally achieve full fault isolability. In the diagnosis system the following sensor configuration is used:

- $\omega_{gb,sens}$ - gearbox outgoing speed
- $U_{b,sens,a}$ - battery voltage
- $U_{b,sens,b}$ - battery voltage

As seen there are two sensors measuring the battery terminal voltage. This is required to be able to achieve full fault isolability, since if only one battery sensor is used, it is not possible to isolate a fault in the battery sensor to all possible faults in the system given by (25a)-(25d).

4 DIAGNOSIS SYSTEMS DESIGN

Two model based diagnosis systems using the sensor sets described in Sections 3.1 and 3.2 are implemented in the platform (Figure 1). The systems should detect

Table 4: Number of equations in the residual generators in Diagnosis system 1 and Diagnosis system 2. The total number of equations is 43 respectively 41 for the two systems.

	# tests	# sensors	# equations
System 1	4	5	2,9,14,12
System 2	6	3	2,15,29,35,35,36

Structurally Overdetermined (MSO) set if there is no subset of \mathcal{M} that is structurally overdetermined (Krysander et al., 2008). There are several efficient tools available to find the MSOs of a model, and some of these tools are discussed and compared in Armengol et al. (2009).

The diagnosis systems designed in this paper are based on MSOs. Based on the vehicle model described in Section 2.3 and the sensor configuration given in Section 3.1, it is possible to construct 65 MSOs, and using the sensor configuration in Section 3.2 it is possible to construct 14 MSOs. These sets of equations are sensitive to different sets of faults, and a selection of the MSOs to be used in the diagnosis systems is made to structurally achieve full fault isolability.

4.1 DIAGNOSIS SYSTEM 1

In the first diagnosis system, five sensors are used as described in Section 3.1. There are several combinations of residual generators that require four residual generators to structurally achieve full fault isolability, and one of these combinations is chosen as the basis to Diagnosis system 1. The residual generators in this diagnosis system are relatively small (see Table 4), as 2-14 equations are used in each residual generator. The sets of equations that are used in the residual generators form substitution chains, and there are no algebraic loops that need to be solved. This leads to that it is easy to design the residual generators.

To illustrate how the residual generators used in the diagnostic tests are constructed, the first residual generator in the diagnosis system is presented. It is based on two equations and one unknown variable U_{em} . This voltage is first computed as

$$U_{em} = U_{em,ctrl} \quad (26)$$

and voltage sensor is then used to compute the residual

$$r = U_{em,sens} - U_{em} \quad (27)$$

and the residual is sensitive for f_{pe} and $f_{em,U,sens}$. The information about which residual generators, or tests, that are expected to react on each fault is summarized in a decision structure in Table 5. Test 1 reacts for example on f_{pe} and $f_{em,U,sens}$, which also can be seen in (25c), (25e), (26) and (27).

Table 5: Decision structure in Diagnosis system 1.

	$f_{em,ka}$	$f_{em,R}$	f_{pe}	$f_{b,sc}$	$f_{em,U,sens}$
$T1$			X		X
$T2$	X	X	X		
$T3$		X		X	
$T4$	X			X	X

If Test 1 alarms this is explained by either f_{pe} or $f_{em,U,sens}$. Here a single fault assumption is made, leading to that if also e.g., Test 2 alarms the only possible explanation for this is that the power electronics is broken since this is the only fault that makes both Test 1 and Test 2 alarm according to Table 5. Full fault isolation is possible since each fault influences different sets of tests (Blanke et al., 2006).

MODEL NOT VALID

In the second test, ω_{em} is computed via the angular speed sensor at the outgoing shaft of the gearbox, $\omega_{gb,sens}$, and the gear ratio, u_{gb} , according to

$$\omega_{em} = u_{gb}\omega_{gb,sens} \quad (28)$$

The noise in $\omega_{gb,sens}$ is amplified with the gear ratio when ω_{em} is computed that later is used in the residual. To get equal test significance for all gears, varying noise levels has to be considered. Here a simple approach is adopted and the test quantity is not updated for gears 1-4, where u_{gb} is large.

CUSUM

Due to sensor noise and model errors, the residuals will be nonzero even in the fault free case. To handle this a standard algorithm called CUSUM (Page, 1954; Gustafsson, 2000) is used. The algorithm is based on that a signal, s , is constructed to have a negative expectation value in the fault free case and a positive expectation value when a fault has occurred. The trend of a cumulative sum, g , of s will then contain information about the status of the monitored system. The test quantity, T , is calculated as

$$s(t) = |r(t)| - \nu \quad (29a)$$

$$g(t+1) = g(t) + s(t) \quad (29b)$$

$$T(t) = g(t) - \min_{0 \leq i < t} g(i) \quad (29c)$$

where ν is an offset that ensures that $E\{s(t)\} < 0$ in the fault free case. The size of ν reflects the model error and noise in the model. The system alarms if $T > J$, where J is a threshold and a design parameter, that is set to avoid false alarms and still achieve fast fault detection.

Table 6: Decision structure in Diagnosis system 2.

	$f_{em,ka}$	$f_{em,R}$	f_{pe}	$f_{b,sc}$	$f_{b,U,sens,a}$	$f_{b,U,sens,b}$	$f_{\omega,gb,sens}$
$T1$					X	X	
$T2$		X	X	X		X	X
$T3$	X	X	X				X
$T4$	X	X		X	X		X
$T5$	X		X	X		X	X
$T6$	X	X	X	X		X	

4.2 DIAGNOSIS SYSTEM 2

The second diagnosis system uses as few sensors as possible to structurally achieve full fault isolability in the system. This choice is made to investigate the impact regarding performance of the system, but also the complexity in designing the system. All three sensors are monitored, resulting in that there are seven fault modes considered in Diagnosis system 2.

To achieve full fault isolability six tests are needed in this diagnosis system compared to four in Diagnosis system 1. It is common that models are not accurate or valid in all operating modes of the system. To investigate how this influence the diagnostic performance, the model of the clutch used in the diagnosis design is only valid when the clutch is fully engaged. The models of the clutch and combustion engine are included in four of the residual generators in Diagnosis system 2 (see Figure 2 for the vehicle configuration). These four residual generators include a differentiated signal, which may lead to problems since the signal is noisy.

The tests in this system are in general based on more model equations than the tests in Diagnosis system 1. Up to 36 equations, of the in total 41 equations describing the truck, are used in the tests (see Table 4). The decision structure for Diagnosis system 2 is given in Table 6.

ALGEBRAIC LOOPS

In five of the six sets of equations, one equation is used to compute one of the unknown variables at the time in the computation sequence of the unknown variables. However, in Test 4 an algebraic loop has to be solved since the following two equations are included in the computation sequence and based on

these two equations I_{em} and U_{em} are to be computed

$$I_{em} = \frac{U_{em} - \omega_{em}k_a}{R_{em}} \quad (30a)$$

$$U_{em} = \frac{I_b U_b}{I_{em}} \quad (30b)$$

These two equations are only a small part of the residual generator, and the entire consistency relation is given in Sundström (2011). The equation system above has two solutions for I_{em}

$$I_{em} = -\frac{\omega_{em}k_a}{2R_{em}} \pm \sqrt{\left(\frac{\omega_{em}k_a}{2R_{em}}\right)^2 + \frac{I_b U_b}{R_{em}}}, \quad (31)$$

and both solutions are valid, but in different operating modes of the electric machine. In this case, the consistency based diagnosis approach is to alarm when none of the solutions are consistent with measurement data. Thus, two residuals are computed and the residual with the lowest magnitude is used in CUSUM to calculate the test quantity. The disadvantage with this is that the computational burden increases compared to compute only one residual per test quantity.

MODEL NOT VALID

The model of the clutch in the diagnosis system is only valid when the clutch is fully engaged. In this operating mode, the model is that the torques and speeds on both sides of the clutch are equal. This results in that when the clutch is disengaged or there is a slip in the clutch, the four tests that include the clutch model are not valid. After the clutch pedal is fully released by the driver, the residuals in the corresponding tests are not updated in 3 seconds.

Tests 5 and 6 are noise sensitive for small U_{em} , so voltages close to zero needs to be handled. Therefore the tests are not updated when $|U_{em}| < 1$ V.

DYNAMIC RESIDUAL GENERATORS

As stated above, four of the six residual generators in Diagnosis system 2 include dynamics. These residual generators use the same consistency relation, $0 = a\dot{\omega}_{gb} + b$ where a is a constant and b is an arbitrary function of known signals, but the signals a , b , and ω_{gb} are computed in different ways in the different residual generators. The differentiated variable, ω_{gb} , is computed from control inputs and sensor signals, but since sensor signals include measurement noise, also ω_{gb} includes noise. Therefore the residuals, \tilde{r} , are filtered to obtain the residual r

$$r = \frac{\alpha}{p + \alpha} \underbrace{(a\dot{\omega}_{gb} + b)}_{\tilde{r}} \quad (32a)$$

where p is the differentiation operator and $\alpha > 0$ for stability.

It is possible to compute r in (32a), without computing a differentiated signal using a variable transformation. Conditions for this to be possible is that the residual generator can be written on the form $\tilde{r} = a\dot{\omega}_{gb} + b$ as above, where a is a constant and b a function of known signals, and the residual is filtered as in (32a) (Frisk and Nyberg, 2001). By introducing the state Γ , the variable transformation is

$$\Gamma = r - \alpha a \omega_{gb} \quad (32b)$$

We obtain that the residual generator in (32a) can be expressed as

$$\dot{\Gamma} = -\alpha \Gamma - \alpha^2 a \omega_{gb} + \alpha b \quad (32c)$$

$$r = \Gamma + \alpha a \omega_{gb} \quad (32d)$$

The filter parameter α can be modified in the design of the low pass filter of the residual, where a smaller α filters the signal more. The disadvantage with this is that if there is an error in the initialization of the signal, it will take longer time before the error has faded out. On the other hand, it may be difficult to detect faults using a faster filter, since the noise in \tilde{r} is more apparent in r in such a case.

When no gear is selected or the clutch is not engaged, the residual as well as the test quantity are not updated as stated above. When the model in the diagnosis system is getting valid, Γ is reinitialized. This is needed since the state will drift during the time the model was invalid. When Γ is initialized, it is assumed that the vehicle is fault free, i.e. $r = 0$ in (32b). The expression for Γ when the model is getting valid at time t_0 is therefore

$$\Gamma(t_0) = -\alpha a \omega_{gb}(t_0) \quad (33)$$

Using this expression in the initialization is not a good idea, since it is sensitive to noise in ω_{gb} at one sample. This can lead to a significant offset in the residual before the transient has faded out. To reduce the problem, the right hand side of (33) is filtered. It is possible to filter the signal since the parts of the model that is used to compute ω_{gb} is valid even when the entire model used in the residual generators is not valid. Figure 6 shows the signal of $\alpha a \omega_{gb}$ for one residual generator in addition to the signal $\alpha a \omega_{gb}$ filtered with different time constants, τ_r . The filter reduces the above described problem in the initialization of the state. The signal is still noisy, but if it is more filtered there are significant errors during transients since a causal filter is used. The chosen time constant for the four dynamic residual generators in this diagnosis system is 0.1 seconds.

To further reduce the issues when reinitializing the state in the residual generator, the CUSUM algorithm is not updated for the first 10 seconds after the model is getting valid. During this time, most of the error in the initialization of Γ will fade out. The drawback is that the test may be inactive a significant part of the time, and thereby reducing the performance of the system.

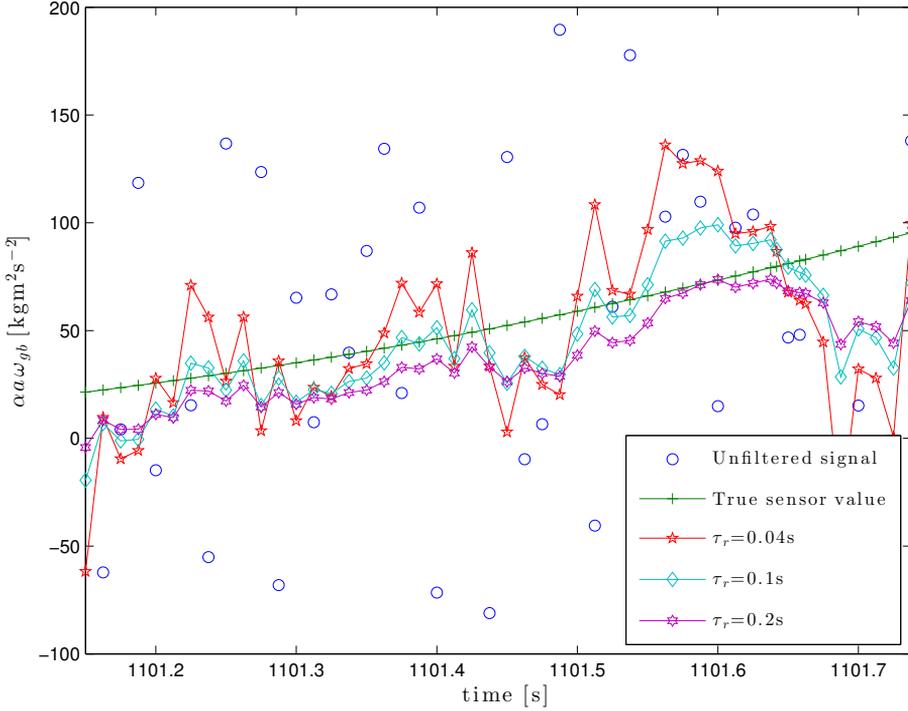


Figure 6: Three time constants in the filter of $\alpha\omega_{gb}$, in addition to the unfiltered signal, are included in the figure. The filter with small time constant is more noisy than the others, while the filter with large time constant has a time delay during the transient. The time constant is set to 0.1 seconds in the diagnosis systems.

5 RESULTS AND DISCUSSION

The designed diagnosis systems achieve full fault isolability according to the structural analysis. To evaluate the performance of the diagnosis systems when noise is added to the sensor signals, simulations of the long haulage truck are carried out. The type of issues handled are e.g. the impact of the number of sensors on the performance of the diagnosis systems, and the interplay between diagnosis and the energy management. These issues are of interest, since they will also occur in reality when developing diagnosis systems. The size of the considered faults are given in Table 3, and the faults are induced one by one in the simulation model to evaluate the diagnostic performance.

The test quantities achieved from the simulations of the diagnosis systems are normalized with the threshold used in the CUSUM algorithm, see Section 4.1

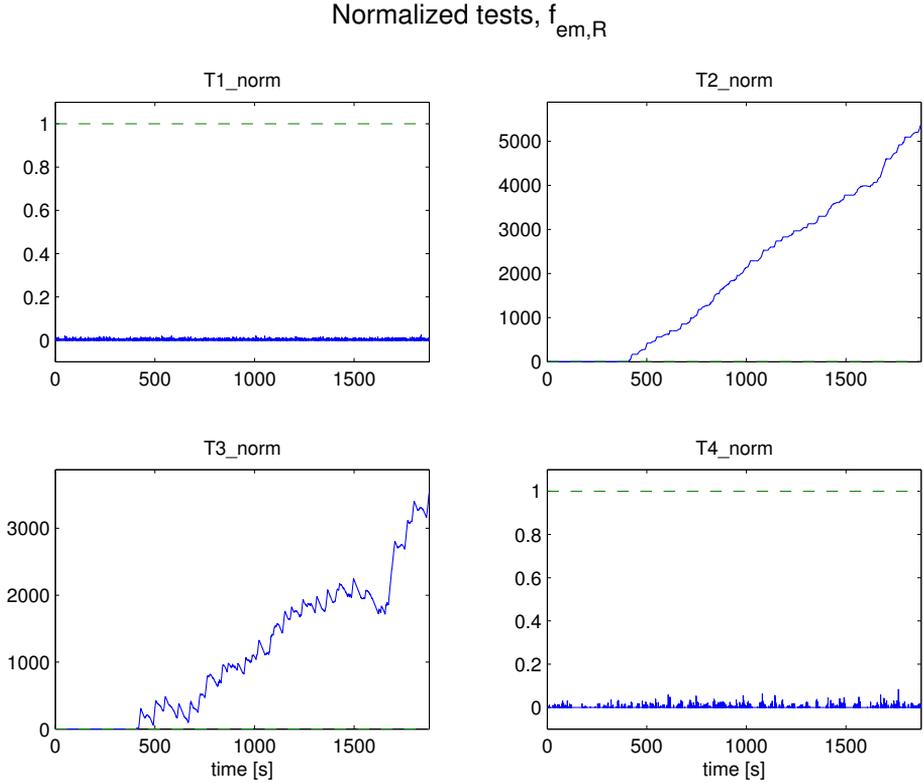


Figure 7: Normalized test quantities in Diagnosis system 1 when $R_{em} = 0.5R_{em}^{nom}$ at 400 seconds and the driving cycle used is FTP75. The tests alarm when the normalized test quantity is larger than one, and Test 2 and Test 3 react on the fault as expected.

for details,

$$T_{norm} = \frac{T}{J} \quad (34)$$

and the test alarms if $T_{norm} > 1$.

5.1 DIAGNOSIS SYSTEM 1

The diagnosis systems based on five sensors detects and isolates all faults in a few seconds. As an illustration, Figure 7 shows T_{norm} when $f_{em,R}$ is induced in the model after 400 seconds and the driving cycle used is FTP75. Test 2 and Test 3 react on this fault, as expected according to the decision structure in Table 5. The performance of the system detecting $f_{em,R}$, is representative for

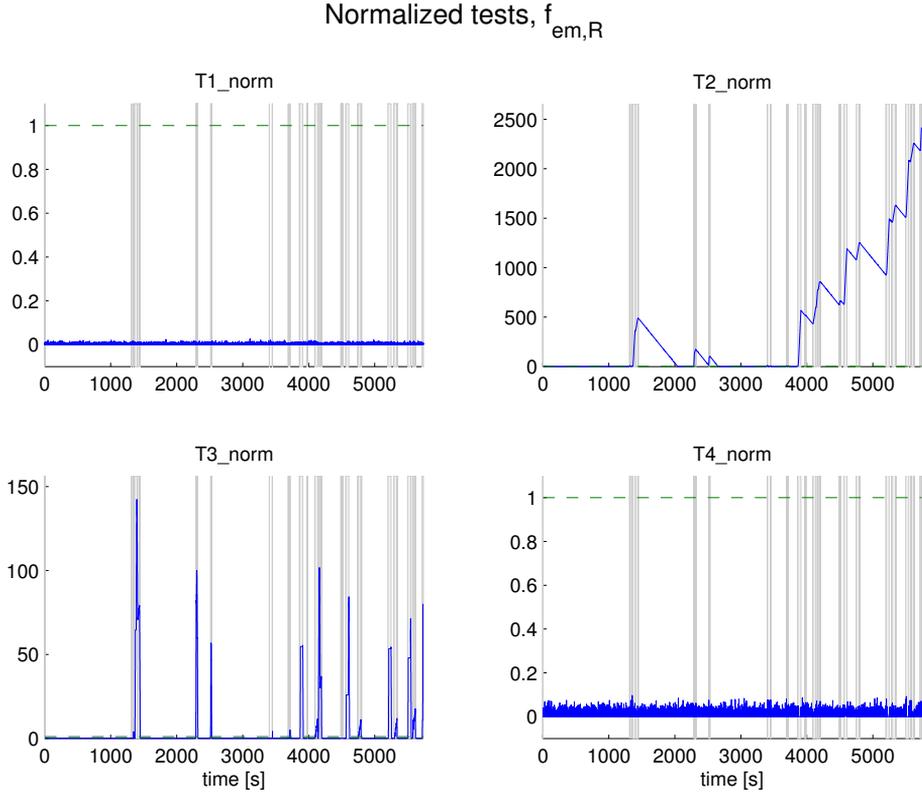


Figure 8: Normalized test quantities in Diagnosis system 1 when $R_{em} = 0.5R_{em}^{nom}$ and the route used is Linköping to Jönköping. The shaded fields indicate when $|I_{em}| > 40A$, and it is clearly shown that the performance in the diagnosis system is dependent on the operating points in the electric machine. Both Test 2 and Test 3 react on the fault as expected.

all faults that are to be detected.

A simulation using the driving profile from Linköping to Jönköping and the resistance in the electric machine is modified after 400 seconds, is carried out and the result is shown in Figure 8. Test 2 and Test 3 react and isolate the fault in this realistic driving scenario. The reason for that the tests do not react at all times on the fault, is that the electric machine is not used during long periods. The shaded areas in the figure indicates when $|I_{em}| > 40A$, which is the current when $|T_{em}| > 20 \text{ Nm}$ for the nominal value of k_a according to (3a). If the test quantities would only be updated when the condition $|I_{em}| > 40A$ is fulfilled, the test quantities in Test 2 and Test 3 would, as expected, be increasing after the fault is induced.

Table 7: Isolability matrix of Diagnosis system 2 based on the results from the simulations of the vehicle model. The fault isolability performance is the same when either of FTP75 or Linköping-Jönköping is used.

	$f_{em,ka}$	$f_{em,R}$	f_{pe}	$f_{b,sc}$	$f_{b,U,sens,a}$	$f_{b,U,sens,b}$	$f_{\omega,gb,sens}$
$f_{em,ka}$	X						X
$f_{em,R}$		X	X				
f_{pe}			X				
$f_{b,sc}$				X			
$f_{b,U,sens,a}$					X		
$f_{b,U,sens,b}$						X	
$f_{\omega,gb,sens}$							X

5.2 DIAGNOSIS SYSTEM 2

As stated above, all faults are fully isolable in Diagnosis system 2 according to the structural analysis. However, the results from the simulation study show that two of the faults are not fully isolable, as can be seen in the isolability matrix in Table 7, where the columns and rows correspond to the faults. An 'X' at position (i, j) indicates that fault i is not isolable from fault j , see e.g., Gelso et al. (2008). In Table 7 it can be seen that the diagnosis system does not isolate $f_{em,ka}$ from $f_{\omega,gb,sens}$, and that $f_{em,R}$ is not isolated from f_{pe} . The reasons are as follows. When the torque constant in the electric machine has changed, i.e. the fault $f_{em,ka}$, Tests 3-5 react, and Test 6 does not react as expected from the structural analysis and the decision structure in Table 6. This means that $f_{em,ka}$ can not be isolated from $f_{gb,\omega,sens}$, since both these faults can be diagnosis statements when Tests 3-5 have reacted, see Table 6. Further, when the resistance in the electric machine has changed, i.e. $f_{em,R}$, Test 4 is not affected as expected. This is the case in both FTP75 and Linköping-Jönköping, and is shown in Figure 9 for FTP75. Due to that Test 4 does not react on the fault, $f_{em,R}$ is not isolated from f_{pe} . Improvements can be sought by using variable parameters in the CUSUM algorithm, that changes with the operating points of the vehicle to adapt to the varying fault sensitivity.

For the five faults that are fully isolable, the result is obtained within 100 seconds. One of the reasons that it takes longer time than for Diagnosis system 1 to reach full fault isolability for these faults is that more of the tests are not valid at all times, here because the model of the clutch is not valid in all operating modes, $|U_{em}|$ is small, or that no gear is selected. A test quantity based on a dynamic residual generator that only is valid when $|U_{em}| > 1$ V, is e.g. updated during 30% of the simulation time when FTP75 is used. In the four tests based on dynamic residual generators, the states in the filters are

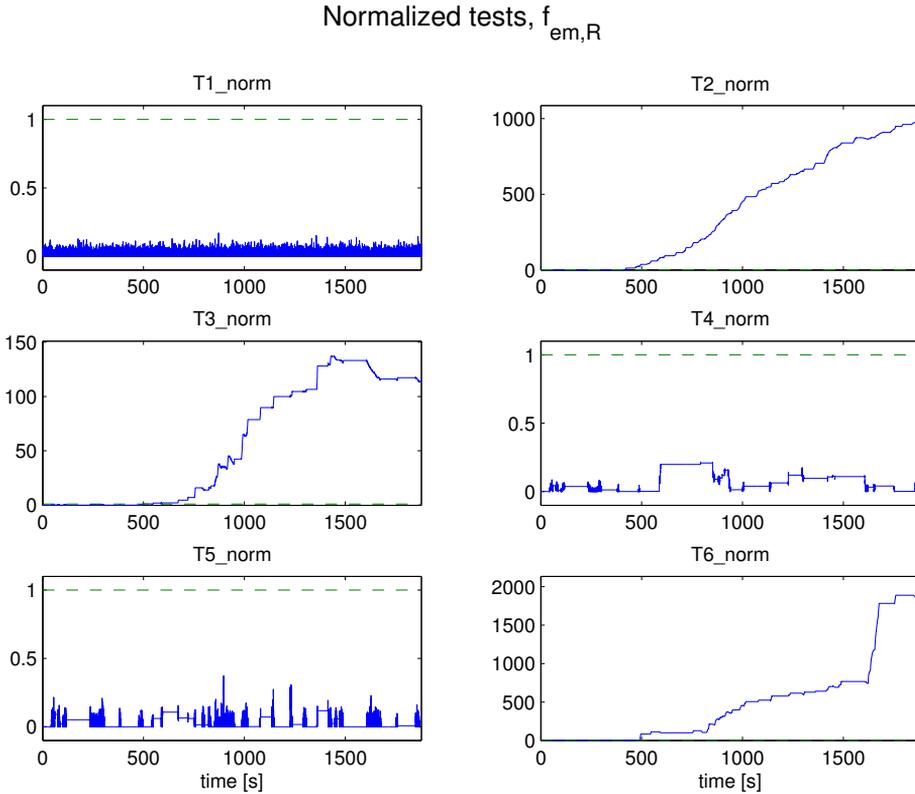


Figure 9: The figure shows the normalized tests when there is a fault in the resistance in the electric machine. Test 4 does not react on the fault as it should do according to the structural analysis.

reinitialized when the system is reactivated. The assumption that the system is fault free is used in the reinitialization of the state. An alternative to this, which possibly increases the diagnostic performance, is to instead use the previous valid value of the residual in the initialization of the state.

6 CONCLUSIONS

The influence of e.g., sensor configuration and operating modes on vehicle level diagnosis has been studied by designing and implementing two diagnosis systems. According to the structural analysis of the model used in these systems, full fault isolability is possible to achieve in both sensor configurations. A simulation study of the implemented diagnosis systems is done and both FTP75 and the realistic

driving scenario Linköping to Jönköping is used. The diagnostic results are similar based on these two driving missions, and the simulation study shows that all faults are fully isolated in the first diagnosis system, that is based on several sensors measuring signals on the components to be monitored. In the second diagnosis system, that is based on a minimal number of sensors to structurally achieve full fault isolability, all faults are not fully isolated in the implemented system, as can be seen in Table 7. The discrepancy between the structural analysis and the performance of the implemented diagnosis system, stems from the influence of the faults on the system in relation to the sensor noise level.

It is shown in Figure 8 that the diagnosis performance is affected by the operating points of the vehicle, which depends on the driving mission and the overall energy management control strategy. This interaction is most significant in the system based on few sensors, and especially in the dynamic residual generators. One main reason for this is that the test quantities are not updated for some time after the model has become valid to reduce the impact of the transient in the reinitialization of the states used in the filters of the residuals. Therefore it is preferable to avoid many deactivations and activations of the tests, and this can be achieved in a well designed energy management.

The overall conclusion is that the performance in the diagnosis system based on several sensors performs better compared to the system based on few sensors. This is an expected result, but if the diagnosis performance is considered when designing the overall energy management, the performance of the latter diagnosis system would be significantly improved. Therefore, by considering the impact of the energy management on the diagnosis system, it may be possible to reduce the number of sensors used in the vehicle to achieve the required diagnostic performance.

REFERENCES

- J. Armengol, A. Bregon, T. Escobet, E. R. Gelso, M. Krysander, M. Nyberg, X. Olive, B. Pulido, and L. Trave-Massuyes. Minimal structurally overdetermined sets for residual generation: A comparison of alternative approaches. In *Proceedings of IFAC Safeprocess'09*, Barcelona, Spain, 2009.
- M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer, 2nd edition, 2006.
- H. A. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, and I. V. Kolmanovsky. Predictive energy management of a power-split hybrid electric vehicle. In *Proceedings of the 2009 conference on American Control Conference*, pages 3970–3976, Piscataway, USA, 2009. IEEE Press.
- R. Bradley. Technology roadmap for the 21st century truck program. Technical Report 21CT-001, U.S. Department of Energy, Oak Ridge, Tennessee, December 2000.
- Z. Chen, Y. Fu, and C. Mi. State of charge estimation of lithium-ion batteries in electric drive vehicles using extended kalman filtering. *Vehicular Technology, IEEE Transactions on*, 62(3):1020–1030, 2013.
- J. de Kleer, A. Mackworth, and R. Reiter. Characterizing diagnoses and systems. *Artificial Intelligence*, 56(2-3):197–222, 1992.
- D. Dustegör, E. Frisk, V. Coquempot, M. Krysander, and M. Staroswiecki. Structural analysis of fault isolability in the DAMADICS benchmark. *Control Engineering Practice*, 14(6):597–608, 2006.
- J. Fredriksson, J. Larsson, J. Sjöberg, and P. Krus. Evaluating hybrid electric and fuel cell vehicles using the capsim simulation environment. In *22nd International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exposition*, pages 1994–2004, Yokohama, Japan, 2006.
- E. Frisk and M. Nyberg. A minimal polynomial basis solution to residual generation for fault diagnosis in linear systems. *Automatica*, 37(9):1417–1424, September 2001.
- E. Gelso, S. Castillo, and J. Amengol. An algorithm based on structural analysis for model-based fault diagnosis. *Artificial Intelligence Research and Development*, 184:138–147, 2008.
- F. Gustafsson. *Adaptive filtering and change detection*. John Wiley & Sons, 2000.
- L. Guzzella and A. Amstutz. CAE tools for quasi-static modeling and optimization of hybrid powertrains. *IEEE Trans. on Vehicular Technology*, 48(6):1762–1769, November 1999.

- L. Guzzella and C. H. Onder. *Introduction to Modeling and Control of Internal Combustion Engine Systems*. Springer Verlag, Zürich, 2004.
- L. Guzzella and A. Sciarretta. *Vehicle Propulsion System, Introduction to Modeling and Optimization*. Springer Verlag, Zürich, 3 edition, 2013.
- I. Husain. *Electric and Hybrid Vehicles*. CRC Press LLC, 2003.
- J. Inhelder. *Verbrauchs- und schadstoffoptimiertes Ottomotor-Aufladekonzept*. PhD thesis, Eidgenössische Technische Hochschule, 1996.
- M. Krysander and E. Frisk. Sensor placement for fault diagnosis. *IEEE Trans. on SMC – Part A*, 38(6):1398–1410, 2008.
- M. Krysander, J. Åslund, and M. Nyberg. An efficient algorithm for finding minimal over-constrained sub-systems for model-based diagnosis. *IEEE Trans. on SMC – Part A*, 38(1), 2008.
- C.-C. Lin, H. Peng, J. Grizzle, and J.-M. Kang. Power management strategy for a parallel hybrid electric truck. *Control Systems Technology, IEEE Transactions on*, 11(6):839 – 849, November 2003.
- E. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.
- T. B. Reddy. *Linden’s Handbook of Batteries*. McGraw-Hill, 2011.
- G. Rizzoni, L. Guzzella, and B. Baumann. Unified modeling of hybrid electric vehicle drivetrains. *Mechatronics, IEEE/ASME Transactions on*, 4(3):246 –257, September 1999.
- A. Sciarretta and L. Guzzella. Control of hybrid electric vehicles. *IEEE Control Systems Magazine*, 27(2):60–70, April 2007.
- M. J. Sivertsson, C. Sundström, and L. Eriksson. Fuel consumption minimization in a two axle hybrid with adaptive ECMS. In *IFAC World Congress*, Milano, Italy, 2011.
- C. Sundström. Vehicle level diagnosis for hybrid powertrains. Technical report, 2011. Licentiate thesis LiU-TEK-LIC-2011:27, Thesis No. 1488.
- L. O. Valøen and M. I. Shoesmith. The effect of PHEV and HEV duty cycles on battery and battery pack performance. In *Proceedings in 2007 Plug-in Highway Electric Vehicle Conference*, Manitoba, Canada, 2007.

A New Electric Machine Model and its
Relevance for Vehicle Level Diagnosis*

B

*Submitted to Journal.

A New Electric Machine Model and its Relevance for Vehicle Level Diagnosis

Christofer Sundström, Lars Nielsen, and Erik Frisk

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, SE-581 83 Linköping, Sweden.*

ABSTRACT

With the electrification of society, especially transportation, the control and supervision of electrical machines become more and more important due to its bearing on energy, environment, and safety. To optimize performance in control and supervision, appropriate modelling is crucial, and this regards both the ability to capture reality and the computational complexity to be useful in real time. Here a new low complexity model of the electric machine is proposed and developed. The new model treats the machine constants in a different way compared to a previous standard model, which results in a different expression for power losses. It is shown that this increases model expressiveness so when adopted to real data the result is significantly better. The significance of this modelling improvement is demonstrated using a task in vehicle diagnosis where it is shown that the separation between the non-faulty and faulty cases is better and the resulting performance is improved.

1 INTRODUCTION

A hybrid vehicle is more complex than a conventional vehicle since it has more components e.g., electric machine, battery, and power electronics (Husain, 2003; Guzzella and Sciarretta, 2013), and it is important to monitor these components due to safety issues and to avoid damage. Following a model based diagnosis approach, the engineering steps are to devise component models and their interconnections, to design residuals and test quantities, to choose thresholds, and finally to state diagnostic decisions (Blanke et al., 2006). In practice, these steps are interlinked and good engineering is needed in each step. Regarding the models, they should of course describe reality sufficiently well, and at the same time be computationally effective to be able to execute the diagnostic system on-board the vehicle.

The main contribution in this paper is a new model for the electric machine. The model has low computational complexity to be able to execute it in real time on-board a vehicle, e.g., in a diagnosis system. Further, the number of parameters in the model is small, which is advantageous in model calibration as well as e.g., in initial studies in powertrain configurations. The proposed model is a modification of a standard model in Guzzella and Sciarretta (2013) keeping the same order of computational complexity. Nevertheless, the principal ability to fit real data is significantly better in the new model, which is demonstrated in Section 2. To demonstrate the value of the new model in vehicle level diagnosis, in Section 4 two diagnosis systems monitoring the electric machine in a hybrid powertrain are designed based on the new model and the standard model respectively. A main result is better separation between the non-faulty and faulty cases in the diagnosis system based on the new model.

2 ELECTRIC MACHINE MODEL

In hybrid electric vehicles (HEV) mainly permanent magnet synchronous machines (PMSM) are used, despite their high costs related to the permanent magnets (Husain, 2003), since this type of machine in general has higher efficiency and power density compared to other machine types (Zhu and Howe, 2007; Chau et al., 2008). Typical efficiency maps for an induction machine and a PMSM are shown in Mellor (1999).

A PMSM consists of a stator with windings, and a rotor with permanent magnets. The magnets are either mounted on the outside of the rotor, or are integrated inside the rotor (Chau et al., 2008). By applying a voltage that results in a current in the stator, the rotor starts to move. A PMSM is an AC machine, but it is possible to use a DC source, e.g., a battery, and use power electronics to achieve an alternating current. The torque generation principle in a PMSM and brushless DC (BLDC) machine is the same (Fitzgerald et al., 2003). The main difference between the two machine types is that the waveform of the stator current is rectangular in the BLDC, but sinusoidal in the PMSM. In a hybrid electric vehicle it is common to use the notation PMSM for both these two types of machines, and therefore this notation is also used here.

Two models of electric machines are presented and evaluated using measurements of the power losses. The first model is a standard model described in Guzzella and Sciarretta (2013), and the second model is a new model that is a modified and extended version of the first model. Both models are static, which is sufficient for the purpose here, but it is straightforward to include, e.g., an inductance and thereby add dynamics to the model.

2.1 STANDARD MODEL

A BLDC or a PMSM can be seen as an inside out DC machine, i.e., with field windings on the rotor and where the stator is electronically commutated using power electronics Fitzgerald et al. (2003). BLDCs are often modelled as separately excited DC machines with constant magnetic field, while PMSMs often are modelled as synchronous machines using the direct and quadrature transformation. The model in this section is recalled from Guzzella and Sciarretta (2013), and the machine is modelled as a separately excited DC machine. The model is denoted as *standard model*, with superscript *std* in some of the variables. The magnetic flux ϕ is constant in a PMSM, and the torque, T_{em}^{std} , is modelled to be proportional to the current I_{em} with the torque constant k_a (Wang et al., 2011; Yildiz, 2012). With K as a machine constant depending on design parameters of the machine, the equations become

$$T_{em}^{std} = k_a I_{em} \quad (1)$$

$$k_a = K\phi. \quad (2)$$

The current in the stator, I_{em} , is calculated using the voltage, U_{em} , supplied by the power electronics, and the electromotive force (emf), that depends on the speed of the machine, ω_{em} , as

$$I_{em} = \frac{1}{R_{em}} (U_{em} - \underbrace{k_i \omega_{em}}_{emf}), \quad (3)$$

where R_{em} is the resistance in the electric machine and k_i the speed constant. Ideally $k_a = k_i$, but here $k_a < k_i$ to model the losses in the machine in addition to the resistive losses. Combining (1) and (3) results in

$$T_{em}^{std} = \frac{k_a}{R_{em}} U_{em} - \frac{k_a k_i}{R_{em}} \omega_{em}. \quad (4)$$

The power losses in the machine are computed by

$$P_{em,l}^{std} = I_{em} U_{em} - T_{em}^{std} \omega_{em}. \quad (5)$$

Substituting U_{em} and I_{em} using (1) and (3) results in the power loss for the standard model as

$$P_{em,l}^{std} = \underbrace{\left(\frac{T_{em}^{std}}{k_a} \right)^2}_{I_{em}^2} R_{em} + \left(\frac{k_i}{k_a} - 1 \right) T_{em}^{std} \omega_{em}. \quad (6)$$

2.2 NEW MODEL

In the *new model*, the current is modelled in the same way, i.e. (3), but in the torque model (1), k_a is modelled differently and losses are explicitly included. The losses in electric machines are resistive losses, friction and windage losses, and iron losses (Udomsuk et al., 2011). The resistive losses are considered in the previous model. Here, in the new model, the friction and windage losses are lumped and are modelled as friction losses. The torque due to friction is modelled to be proportional to ω_{em} (Zhu et al., 2000) by the friction constant $c_{em,f}$ as

$$T_f = c_{em,f} \omega_{em}. \quad (7)$$

The output torque is computed similar to (1), but also considering T_f as

$$T_{em}^{new} = k_a I_{em} - c_{em,f} \omega_{em}. \quad (8)$$

Substituting the current with the expression in (3), which is the same for both models, gives

$$T_{em}^{new} = k_a \left(\frac{U_{em}}{R_{em}} - \frac{k_i}{R_{em}} \omega_{em} \right) - c_{em,f} \omega_{em}. \quad (9)$$

In the standard model all losses are described as resistive losses and by $k_a < k_i$, see (6). In the new model the friction and windage losses are considered in (9), and the resistance is included in the model. The iron losses, P_{Fe} , are included in the new model by using different values for the parameters k_a and k_i . The iron losses can be separated in hysteresis losses, P_h , and eddy-current losses, P_e , and are commonly modelled (Mi et al., 2003) as

$$P_{Fe} = P_e + P_h = k_h B^\beta \omega_s + k_e B^2 \omega_s^2 \quad (10)$$

where k_h and k_e are constants, β the Steinmetz constant that often is a value between 1.8 and 2.2, and B the magnetic field that varies with the angular speed ω_s . It is assumed that the magnetic material in the stator is unsaturated, resulting in that B can be modelled to be proportional to I_{em} . This assumption in combination with (10) results in that the delivered torque by the machine is smaller than what the torque would be without considering the iron losses of the machine, see the schematic illustration in Figure 1. To achieve this characteristics of T_{em} an efficiency $\eta_{em,0} < 1$ is used in the new model as

$$k_a = \begin{cases} k_i \eta_{em,0}, & I_{em} \geq 0 \text{ A} \\ \frac{k_i}{\eta_{em,0}}, & I_{em} < 0 \text{ A}. \end{cases} \quad (11)$$

In this equation $k_a < k_i$ in motor mode and $k_a > k_i$ in generator mode, i.e. $T_{em} < 0$. As mentioned above, losses are in the standard model described by $k_a < k_i$, and using this leads to the non-physical result of a curve in Figure 1 that is above the ideal (dashed) curve in generator mode.

The power losses for the new model are finally computed as in (5), where U_{em} and I_{em} are found by using (3) and (8) respectively.

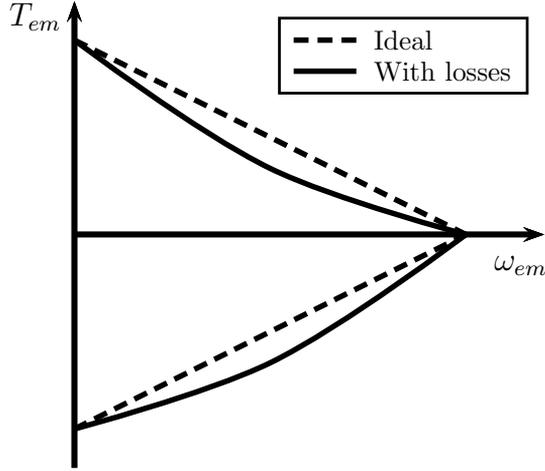


Figure 1: Schematic illustration of the torque delivered by the machine ideally (dashed lines) and when considering the iron losses (solid lines) given some currents and voltages.

$$P_{em,l}^{new} = I_{em}U_{em} - T_{em}^{new}\omega_{em} \quad (12a)$$

$$\begin{aligned} &= (T_{em}^{new} + c_{em,f}\omega_{em})\frac{1}{k_a}(I_{em}R_{em} + k_i\omega_{em}) - T_{em}^{new}\omega_{em} \\ &= \frac{R_{em}}{k_a^2}(T_{em}^{new})^2 + \left(\frac{c_{em,f}^2R_{em}}{k_a^2} + \frac{k_i}{k_a}c_{em,f}\right)\omega_{em}^2 + \\ &\quad + \left(2\frac{c_{em,f}R_{em}}{k_a^2} + \left(\frac{k_i}{k_a} - 1\right)\right)T_{em}^{new}\omega_{em} \end{aligned} \quad (12b)$$

2.3 PARAMETRIZATION OF THE MODELS

The parameters in the models described in Sections 2.1 and 2.2 are identified using measurement data of the power losses of a 59 kW PMSM included in an electric rear axle hybrid vehicle. The map describing the losses, $P_{em,l}^{map}$, takes T_{em} , ω_{em} , and the battery voltage, U_b , as inputs, leading to

$$P_{em,l}^{map} = f(T_{em}, \omega_{em}, U_b). \quad (13)$$

The efficiency map of the machine is given in Figure 2.

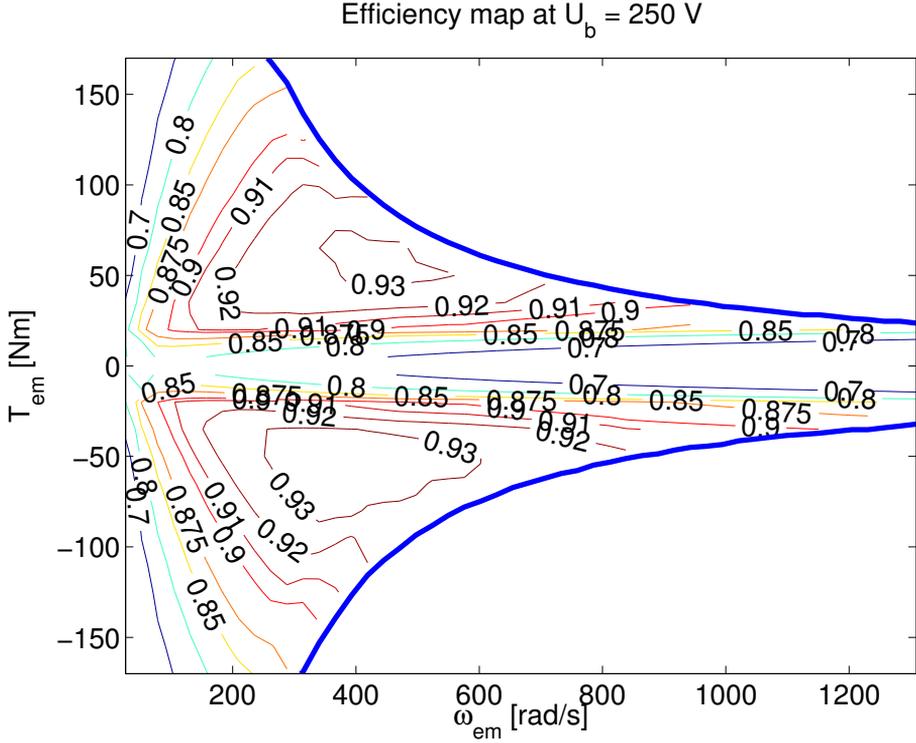


Figure 2: The efficiency of the permanent magnet synchronous machine based on measurements for $U_b=250$ V, and the thick blue lines show the torque limitation of the machine.

PARAMETRIZATION OF THE STANDARD MODEL

For the standard model given in Section 2.1, the power losses of the expression given in (6) are to be fitted to the measured losses given in (13). There are three parameters to be identified, k_a , k_i , and R_{em} . However, these are only included in two terms in the expression (6), leading to that all parameters cannot be identified. Using

$$k_{em,1} = \frac{R_{em}}{k_a^2} \quad (14)$$

$$k_{em,2} = \frac{k_i}{k_a} \quad (15)$$

instead gives

$$P_{em,l}^{std} = (T_{em}^{std})^2 k_{em,1} + (k_{em,2} - 1) T_{em}^{std} \omega_{em}. \quad (16)$$

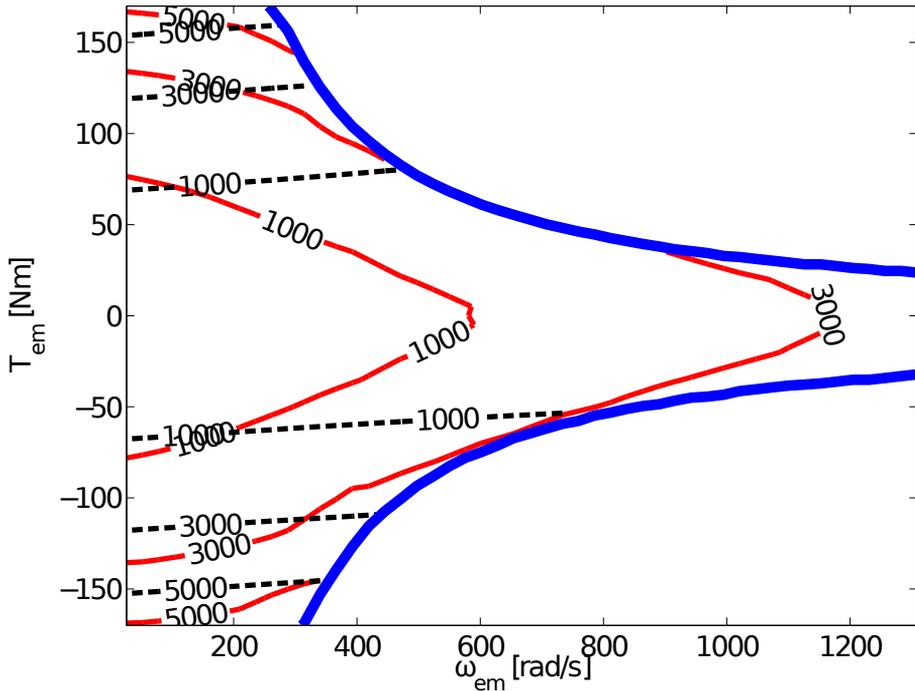


Figure 3: The power losses [W] of the electric machine. The dashed lines are the parametrized model described in Section 2.1, and the solid lines the losses in the map.

The battery voltage is not included in (16), but is required in the map based model. In the parametrization of the model, the battery voltage is assumed to its open circuit voltage, i.e. 250 V. The values of the parameters $k_{em,1}$ and $k_{em,2}$ in (16) are identified by minimizing the squared error between $P_{em,l}^{std}$ in (16) and the data from the map. The values of the parameters found are $k_{em,1} = 0.27 \text{ } \Omega\text{A}^2/\text{N}^2\text{m}^2$ and $k_{em,2} = 0.99$, and the power losses in the electric machine in the map and the parametrized equation (16) are shown in Figure 3. The characteristics of the losses are not captured in the model, since the contour lines of the modelled losses (dashed lines) are almost straight lines while the contour lines of the measured losses (solid lines) have a peak at $T_{em} = 0 \text{ Nm}$. Further, since the estimated value of $k_{em,2} < 1$ then $k_a > k_i$ and not $k_a < k_i$ as expected by (Guzzella and Sciarretta, 2013).

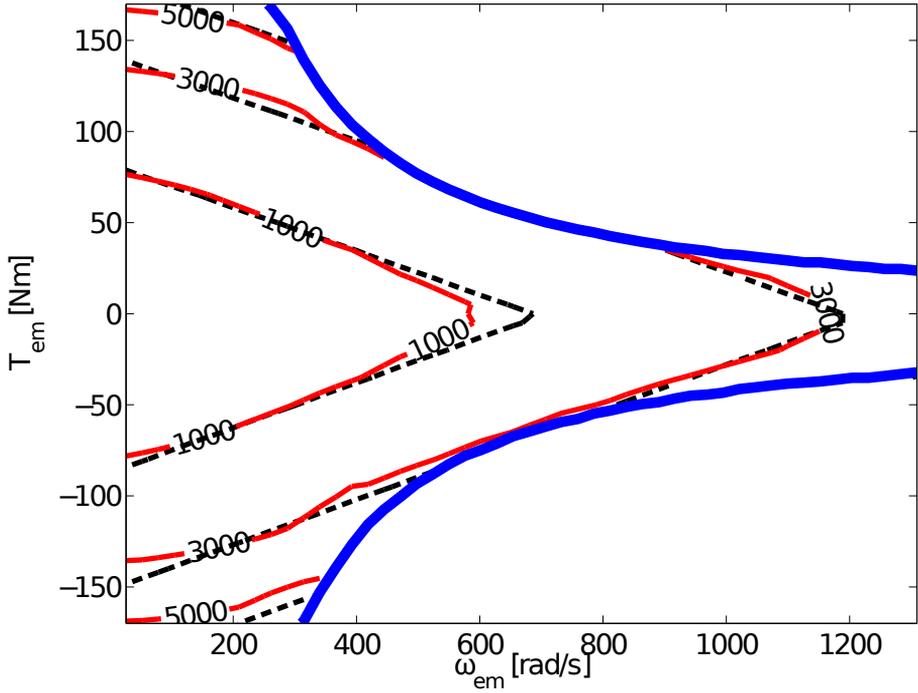


Figure 4: The power losses [W] of the electric machine. The dashed lines are the parametrized model described in Section 2.2, and the solid lines the losses in the map.

PARAMETRIZATION OF THE NEW MODEL

Values for the parameters used in the new model are found by finding the least square error for $P_{em,l}^{new}$ in (12b) using the measured data. The values for k_i , $\eta_{em,0}$, R_{em} , and $c_{em,f}$ are found to be 0.498 Nm/A , 0.97 , $0.039 \text{ } \Omega$, 0.0021 Nm/s , respectively. The power losses computed in (12b) for these parameter values are shown in Figure 4, where also the measured data is shown. As seen in Figure 4, (12b) captures the characteristics of the losses, which was not the case in the standard model, and the modelled losses accurately fit the measured losses.

3 VEHICLE MODEL

The use of the new electric machine model in the design of a diagnosis system is illustrated in Section 4, and the application is the powertrain of a long haulage hybrid electric truck. It is important to evaluate the diagnosis system of a component in the powertrain under realistic operating speeds and torques, and

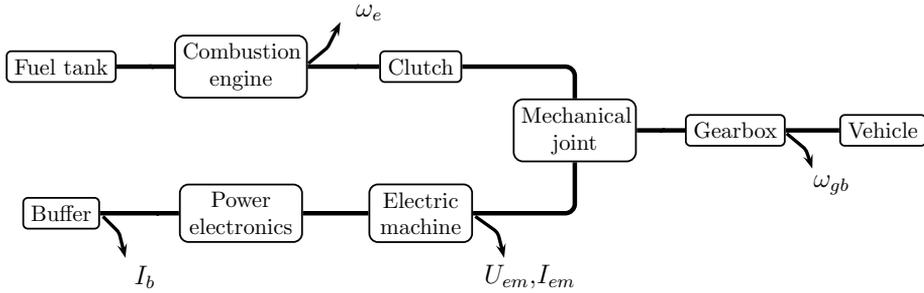


Figure 5: The configuration of the modelled vehicle.

therefore a vehicle model is used in combination with a driving cycle. The powertrain model used in the evaluation is described in this section and is a mean value model suitable for vehicle level diagnosis.

3.1 POWERTRAIN MODEL

The connection of the electrical and conventional parts of the powertrain is located between the clutch and the gearbox, see Figure 5. The model of the powertrain, except for the electric machine, is described in Sundström et al. (2014), where also the model equations are given. The component models of interest for the analysis to come in the next section is recalled below. The vehicle model is implemented in MATLAB/SIMULINK and is based on the model libraries developed in the Center for Automotive Propulsion Simulation (Fredriksson et al., 2006) and the QSS toolbox (Guzzella and Amstutz, 1999). Modifications to these models are carried out to model a truck instead of a passenger car, and to include the possibility to induce faults in the models as well as to add sensor noise. The energy management is based on a heuristic control strategy, and the basic idea is to charge the battery when the vehicle is braking and use this energy as soon as the vehicle requires tractive power (see Sundström et al. (2010) for details). The internal combustion engine is modelled using a Willans line (Guzzella and Sciarretta, 2013; Rizzoni et al., 1999).

BATTERY

The battery is modelled using a Thévenin equivalence circuit, i.e. a voltage source, U_{oc} , that in this case depends on the state of charge, SoC , and an inner resistance, R_b , connected in series (Reddy, 2011). The battery voltage, U_b , on the terminals for a current I_b is

$$U_b = U_{oc}(SoC) - R_b I_b \quad (17)$$

and the state of charge is based on I_b and the capacity of the battery, Q_b , is computed by

$$SoC = SoC_0 - \frac{1}{Q_b} \int I_b dt, \quad SoC \in [0, 1], \quad (18)$$

where SoC_0 is the initial state of charge of the battery.

The modelled battery pack is based on 200 cells, and the cells used are denoted as the “10-Mile PHEV” in Nelson et al. (2007). For the battery pack the resistance is assumed to be a constant, the $U_{oc}(SoC = 0.5)$ is 250 V, and the capacity is 33.6 Ah, or approximately 8.3 kWh.

POWER ELECTRONICS

The efficiency in the power electronics, η_{pe} is assumed to be constant, thus

$$P_b = \eta_{pe}^{-\text{sign}\{P_{em,e}\}} P_{em,e} \quad (19)$$

where P_b and $P_{em,e}$ are the electrical powers from the battery and machine. Using $P_b = I_b U_b$ gives the following expression for the battery current

$$I_b = \eta_{pe}^{-\text{sign}\{P_{em,e}\}} \frac{P_{em,e}}{U_b}. \quad (20)$$

ELECTRIC MACHINE

The map based model used in the parametrization of the electric machine in Section 2.3 is used in the simulation model of the truck. Note that the two models presented in Sections 2.1 and 2.2 are used in the design of residuals to be used in a diagnosis system (see Section 4). The map based model is based on the map describing the power losses of the machine, $P_{em,l}^{map}$, described in (13). The requested torque, that is equal to the delivered torque as long as the machine can deliver the torque, the speed of the machine, and the battery voltage are input signals to the model, see Figure 6. The mechanical power delivered by the machine is expressed as

$$P_{em,m} = T_{em} \omega_{em} \quad (21)$$

and is used to calculate the electrical power as

$$P_{em,e} = P_{em,m} + P_{em,l}^{map}, \quad (22)$$

that is used to compute I_b by using (19).

VEHICLE

In the vehicle, that has a mass m_v and a lumped inertia J_{tot} , the output shaft from the gearbox is connected to the final gear and finally to the wheels. The gear ratio in the final gear is denoted γ_f , and the angular acceleration of the

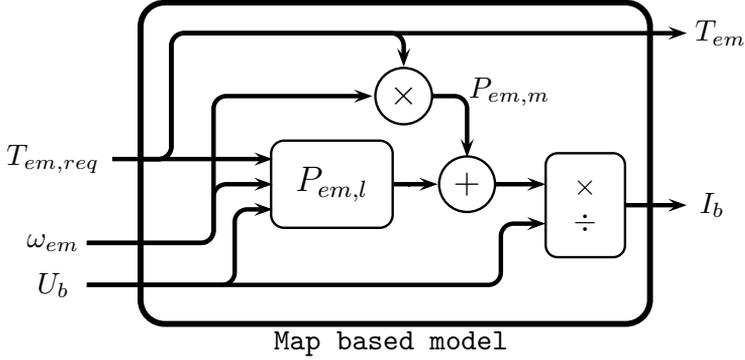


Figure 6: Schematic overview of the signal flow in the map based model, that also includes the model for the power electronics described in Section 3.1.

wheels is calculated based on the resulting torque, T_{net} , acting on the wheels, whose radius is r_w , by

$$T_{net} = T_{gb}\gamma_f - T_d - T_r - T_B - T_g \quad (23a)$$

$$\dot{\omega}_w = \frac{T_{net}}{J_{tot}\gamma_f^2 + m_v r_w^2}, \quad (23b)$$

where T_{gb} , T_d , T_r , T_B , and T_g are the torques from the gearbox, air drag, rolling resistance, mechanical brakes, and slope of the road, respectively. The input torque to the gearbox is the sum of the torques from the electric machine and the clutch, i.e. the internal combustion engine torque when the clutch is engaged (see Figure 5).

3.2 SENSORS

Several sensors are used for control of the vehicle, but only T_{em} and ω_{em} are used in the residuals used in the diagnosis systems presented in Section 4. A measurement signal, y , is given by the noise free signal, y^* , added with noise, \tilde{v}

$$y = y^* + \tilde{v}. \quad (24)$$

If a torque sensor is not available it is possible to use other sensors and an observer as a virtual torque sensor. Here, a torque sensor is used to demonstrate the general principle behavior since the main contribution in this paper is the new electric machine model, and the vehicle model is used to quantify the diagnostic improvement of the increased model performance in a realistic operating scenario.

4 EVALUATION ON DIAGNOSIS SYSTEM

The diagnostic benefit of the new model is illustrated in this section by using the simulation model of the hybrid truck described in Section 3. Two residuals, r_{std} and r_{new} , that should be zero in a fault free case and non-zero in a faulty case, are compared. The first residual is based on the standard model of the electric machine, and the latter is based on the new model.

4.1 INDUCED FAULT

The faults induced in this evaluation are changes in the inner resistance R_{em} and the speed constant k_i of the machine expressed by

$$f_{em,R} : R_{em} = (1 + f_{em,R}) R_{em}^{nom} \quad (25a)$$

$$f_{em,ki} : k_i = (1 + f_{em,ki}) k_i^{nom}, \quad (25b)$$

where R_{em}^{nom} and k_i^{nom} are the nominal values of the resistance and the speed constant, and $f_{em,R}$ and $f_{em,ki}$ the sizes of the corresponding faults. In the map based model that is used to represent the electric machine in the simulations, R_{em} and k_i are not explicitly included. Therefore the map of the power losses and the delivered torque of the machine are modified when there is a fault in the machine, as described in Sundström et al. (2013). In brief, this means that the nominal values for T_{em} and $P_{em,l}$ are given by the map based model, and the modifications due to the faults in (25) are computed using the new model. It is important to note that the map based model is thus assumed to be the truth, and this model is not available in the diagnosis system design.

4.2 RESIDUALS

In the residuals, the signal from a torque sensor in the electric machine, $T_{em,sens}$ is compared to a computed torque, T_{em}^{std} and T_{em}^{new} , respectively

$$r_{std} = T_{em,sens} - T_{em}^{std} \quad (26a)$$

$$r_{new} = T_{em,sens} - T_{em}^{new}. \quad (26b)$$

The computed torque in r_{new} is based on the model described in Section 2.2 and is calculated using the substitution chain given in (27). The gear ratios in the gearbox and the mechanical joint are γ_{gb} and γ_{em} respectively, the output angular speeds from the mechanical joint and gearbox are denoted as ω_{mj} and ω_{gb} respectively, and the requested voltage from the power electronics is denoted

as $U_{em,ctrl}$. The substitution chain used in r_{new} is

$$\begin{aligned}
 \omega_{gb} &= \omega_{gb,sens} \\
 \gamma_{gb} &= f(\text{gear}) \\
 \omega_{mj} &= \gamma_{gb}\omega_{gb} \\
 \omega_{em} &= \gamma_{em}\omega_{mj} \\
 U_{em} &= U_{em,ctrl} \\
 k_a &= \begin{cases} k_i\eta_{em,0}, & T_{em,sens} + c_{em,f}\omega_{em} \geq 0 \\ \frac{k_i}{\eta_{em,0}}, & T_{em,sens} + c_{em,f}\omega_{em} < 0 \end{cases} \quad (27a)
 \end{aligned}$$

$$T_{em}^{new} = k_a \left(\frac{U_{em}}{R_{em}} - \frac{k_i}{R_{em}}\omega_{em} \right) - c_{em,f}\omega_{em}. \quad (27b)$$

The modification in the substitution chain used in r_{std} is that k_a is set to a constant instead of (27a), and that (27b) is replaced by (4), see Section 2.1.

An alarm is generated when the absolute value of a residual is above a specified threshold, J , as

$$\begin{aligned}
 |r| \geq J &\Rightarrow \text{Alarm} \\
 |r| < J &\Rightarrow \text{No alarm.}
 \end{aligned}$$

Ideally, the residual r is zero in the fault-free case but due to noise and model uncertainties the threshold J has to be selected to obtain a suitable trade-off between probability for detection $P(\text{detection})$ and probability for false alarm $P(\text{false alarm})$. A small J results in a high false alarm rate, while a high J results in a low fault detection rate.

4.3 SIMULATION RESULTS

The residuals are post processed using a low pass filter to reduce the noise level. The driving cycle used in the simulations is FTP75 which is 1877 seconds long, i.e., defined on $\tau = [0, 1877]$. The errors are induced one by one in the vehicle model, are constant in the entire simulations, and are set to $f_{em,R} = -0.6$ and $f_{em,ki} = -0.03$ respectively, corresponding to similar errors in torque.

ILLUSTRATION OF SEPARATION PROPERTIES

Figures 7 and 8 zoom in on the time interval [1633, 1639], where r_{std} and r_{new} are plotted in the non-faulty case and for the faulty cases $f_{em,R} = -0.6$ and $f_{em,ki} = -0.03$ respectively. Some interesting differences in separation properties between r_{std} and r_{new} are found. In Figure 7 it can be seen in the upper plot that r_{new} is close to zero in the fault free case, which is not the case in r_{std} seen in the middle plot. The reason for that r_{std} is non-zero is that the modelled torque T_{em}^{std} does not accurately represent the delivered torque of the machine, see (26a). The figure also shows that the separation is larger in r_{new} compared

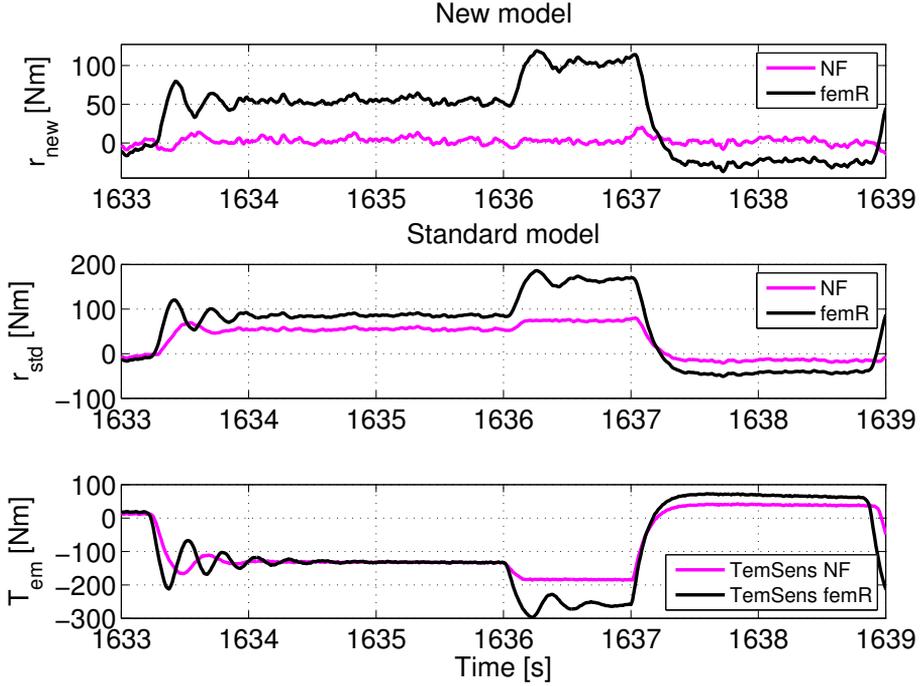


Figure 7: Residuals when $R_{em} = 0.4R_{em}^{nom}$. There is a larger separation between the residuals in the faulty and fault free cases in the new model. The value of the residual based on the standard model in the fault free case, is dependent on the delivered torque T_{em} .

to r_{std} , see e.g., at time 1635 seconds. Relating the lower plot to the upper two plots, it can be seen that the residuals depend on the delivered torque from the machine, but r_{new} is close to zero in the fault free case independent of T_{em} . Also in Figure 8 r_{new} is close to zero independent of T_{em} in the fault free case. However, compared to when $f_{em,R}$ has occurred, r_{new} when $f_{em,ki}$ has occurred is less dependent on the operating point of the vehicle. Comparing the lower plots in Figure 7 and Figure 8 it is found that the delivered torque from the machine is more affected by $f_{em,ki}$ than $f_{em,R}$.

DETECTION PERFORMANCE

In this section the detection performance of the residuals in the entire driving cycle is analyzed. The analysis is presented for $f_{em,R}$, and the results are similar for $f_{em,ki}$. In Figure 7 it is clear that the fault sensitivity in the residual depends on the torque level in that a higher torque gives a clearer fault response in both residuals. Therefore, the residuals are only computed when the torque is above a

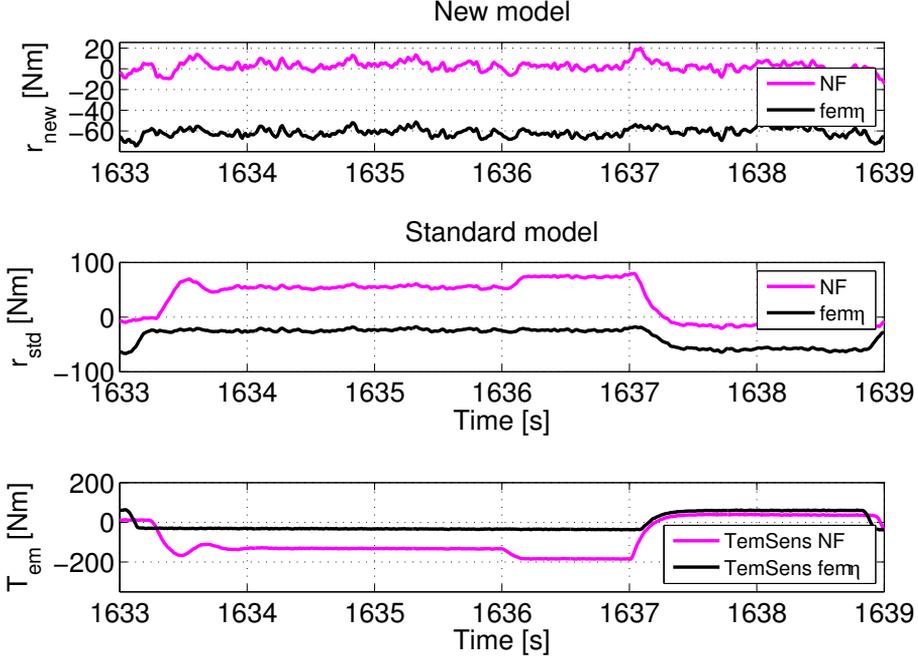


Figure 8: Residuals and delivered torque from the machine when $k_i = 0.97k_i^{nom}$.

threshold, and simulations are carried out for many different torque levels used for residual activation to detect a modification in the resistance of the machine. For illustration, the results based on two of these torque levels, $\beta = 50, 80$ Nm, are included in the paper, which leads to the active time intervals

$$\tau_\beta = \{t \in \tau : |T_{em}(t)| \geq \beta\}, \quad \beta = 50, 80 \text{ Nm}. \quad (28)$$

Now, to study the detection performance for the two residuals, the receiver operating characteristic (ROC) curve is used (Kay, 1998). The ROC curve is the probability for detection, $P(\text{detection})$, plotted against the probability of false alarm, $P(\text{false alarm})$. The curve is parameterized by different thresholds, i.e., each point on the curve corresponds to a specific threshold J . Since it is wanted to achieve a low false alarm rate and a high fault detection rate, a curve that is close to the upper left corner is wanted. Figure 9 shows the ROC curves for the two cases $\beta = 50, 80$ Nm, upper and lower plot. In both cases, it is seen that the solid line, corresponding to the residual r_{new} , is more to the upper left corner than the dashed line, corresponding to the residual r_{std} . Thus, the performance of r_{new} is better than the performance of r_{std} , since the detection rate is higher for the same false alarm rate. The same result is obtained for any value of the parameter β .

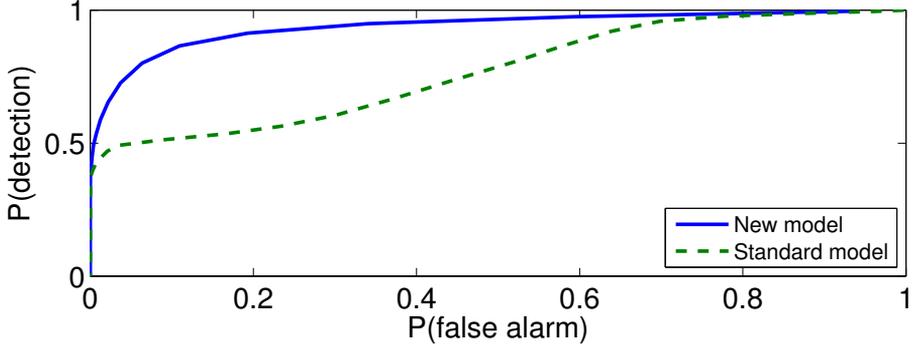
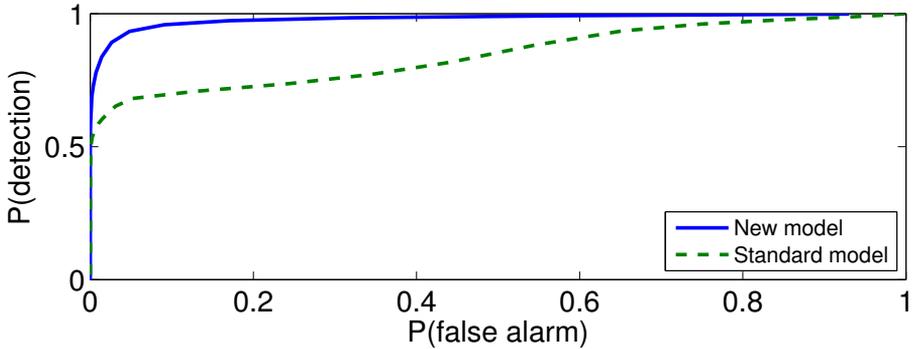
(a) $\beta = 50 \text{ Nm}$ (b) $\beta = 80 \text{ Nm}$

Figure 9: ROC functions for the residuals in Figure 7, where $f_{em,R}$ is considered, for two different values of β in (28). The curves show that the performance is better for the new model, and this holds for any value of β .

5 CONCLUSIONS

The diagnosis task for hybrid vehicles is complex with several interlinked aspects. Here it has been demonstrated that a new low complexity model of the electric machine better describes the component itself and, furthermore, improves diagnostic performance. Compared to the standard model, the new model treats the machine constants in a different way. With this modification, the expression for the power losses becomes different, see (6) and (12b), and these different expressions have different principle capability to model real power losses, as clearly demonstrated in Figures 3 and 4. To investigate the significance of this modelling improvement, a task in vehicle diagnosis is presented and evaluated. The conclusion from Figures 7-9 is that the new model leads to more separated residual signals in the faulty and non-faulty cases compared to the standard

model, as well as less dependence on the operating point of the machine. In conclusion, without increasing complexity, a new useful model of an electric machine has been presented.

REFERENCES

- M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer, 2nd edition, 2006.
- K. Chau, C. Chan, and C. Liu. Overview of permanent-magnet brushless drives for electric and hybrid electric vehicles. *Industrial Electronics, IEEE Trans. on*, 55(6):2246–2257, June 2008.
- A. Fitzgerald, J. Charles Kingsley, and S. D. Umans. *Electric Machinery*. McGraw - Hill, Singapore, 6th edition, 2003.
- J. Fredriksson, J. Larsson, J. Sjöberg, and P. Krus. Evaluating hybrid electric and fuel cell vehicles using the capsim simulation environment. In *22nd International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exposition*, pages 1994–2004, Yokohama, Japan, 2006.
- L. Guzzella and A. Amstutz. CAE tools for quasi-static modeling and optimization of hybrid powertrains. *IEEE Trans. on Vehicular Technology*, 48(6):1762–1769, November 1999.
- L. Guzzella and A. Sciarretta. *Vehicle Propulsion System, Introduction to Modeling and Optimization*. Springer Verlag, Zürich, 3 edition, 2013.
- I. Husain. *Electric and Hybrid Vehicles*. CRC Press LLC, 2003.
- S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998.
- P. Mellor. High efficiency drive-trains for electric and hybrid vehicles. In *Electrical Machine Design for All-Electric and Hybrid-Electric Vehicles (Ref. No. 1999/196)*, *IEE Colloquium on*, pages 8/1–8/7, 1999.
- C. Mi, G. Slemon, and R. Bonert. Modeling of iron losses of permanent-magnet synchronous motors. *Industry Applications, IEEE Transactions on*, 39(3):734–742, May 2003.
- P. Nelson, K. Amine, A. Rousseau, and H. Yomoto. Advanced lithium-ion batteries for plug-in hybrid-electric vehicles. In *Proceedings in 23rd International Electric Vehicle Symposium*, pages 1589–1605, Anaheim, CA, 2007.
- T. B. Reddy. *Linden's Handbook of Batteries*. McGraw-Hill, 2011.
- G. Rizzoni, L. Guzzella, and B. Baumann. Unified modeling of hybrid electric vehicle drivetrains. *Mechatronics, IEEE/ASME Transactions on*, 4(3):246–257, September 1999.
- C. Sundström, E. Frisk, and L. Nielsen. Overall monitoring and diagnosis for hybrid vehicle powertrains. In *6th IFAC Symposium on Advances in Automotive Control*, pages 119–124, Munich, Germany, 2010.

- C. Sundström, E. Frisk, and L. Nielsen. Fault monitoring of the electric machine in a hybrid vehicle. In *7th IFAC Symposium on Advances in Automotive Control*, pages 548–553, Tokyo, Japan, 2013.
- C. Sundström, E. Frisk, and L. Nielsen. Selecting and utilizing sequential residual generators in FDI applied to hybrid vehicles. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, 44(2):172–185, February 2014.
- S. Udomsuk, T. Areerak, K. Areerak, and K. Areerak. Power loss identification of separately excited dc motor using adaptive tabu search. *European Journal of Scientific Research*, 60(4):488 – 497, 2011.
- R. Wang, Y. Chen, D. Feng, X. Huang, and J. Wang. Development and performance characterization of an electric ground vehicle with independently actuated in-wheel motors. *Journal of Power Sources*, 196(8):3962 – 3971, 2011.
- A. B. Yildiz. Electrical equivalent circuit based modeling and analysis of direct current motors. *International Journal of Electrical Power & Energy Systems*, 43(1):1043 – 1047, 2012.
- G. Zhu, L.-A. Dessaint, O. Akhrif, and A. Kaddouri. Speed tracking control of a permanent-magnet synchronous motor with state and load torque observer. *Industrial Electronics, IEEE Transactions on*, 47(2):346 –355, 2000.
- Z. Zhu and D. Howe. Electrical machines and drives for electric, hybrid, and fuel cell vehicles. *Proceedings of the IEEE*, 95(4):746 –765, April 2007.

A NOTATION

$c_{em,f}$	Friction constant
k_a	Torque constant
$k_{em,1}$	Used in parametrization of standard model
$k_{em,2}$	Used in parametrization of standard model
k_i	Speed constant
$f_{em,ki}$	Fault in the machine (k_i modified)
$f_{em,R}$	Fault in the machine (R_{em} modified)
m_v	Vehicle mass
r_{new}	Residual based on the new model
r_{std}	Residual based on the standard model
I_b	Battery current
I_{em}	Electric machine current
J	Threshold for decision making of residuals
J_{tot}	Lumped vehicle inertia
K	Machine constant
P_b	Battery power
$P_{em,e}$	Electric machine power (electrical)
$P_{em,m}$	Electric machine power (mechanical)
$P_{em,l}^{map}$	Power losses in the map
$P_{em,l}^{new}$	Power losses in the new model
$P_{em,l}^{std}$	Power losses in the standard model
Q_b	Battery capacity
R_b	Battery resistance
R_{em}	Electric machine resistance
SoC	State of charge
SoC_0	Initial state of charge
T_d	Torque due to air drag
T_{em}^{new}	Delivered torque in the new model
T_{em}^{std}	Delivered torque in the standard model
$T_{em,sens}$	Torque sensor in electric machine
T_g	Torque due to road slope
T_{gb}	Torque from gearbox
T_{net}	Resulting torque acting on the wheels
T_r	Torque due to rolling resistance
T_f	Friction torque
T_B	Brake torque
U_b	Battery terminal voltage
U_{em}	Electric machine voltage
$U_{em,ctrl}$	Requested voltage from machine controller
U_{oc}	Battery open circuit voltage
β	Residual activation threshold

γ_{em}	Mechanical joint gear ratio
γ_f	Final drive gear ratio
γ_{gb}	Gearbox gear ratio
$\eta_{em,0}$	Efficiency to model iron losses
η_{pe}	Power electronics efficiency
ϕ	Magnetic flux
ω_{em}	Electric machine speed
ω_{gb}	Gearbox output angular speed
ω_{mj}	Mechanical joint output angular speed

Selecting and Utilizing Sequential Residual
Generators in FDI Applied to Hybrid Vehicles*

C

*Published in *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 44(2):172-185, 2014.

Selecting and Utilizing Sequential Residual Generators in FDI Applied to Hybrid Vehicles

Christofer Sundström, Erik Frisk, and Lars Nielsen

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, SE-581 83 Linköping, Sweden.*

ABSTRACT

For a realistic model of a complex system there will be thousands of possible residual generators to be used for diagnosis. Based on engineering insights of the system to be monitored, certain algebraic and dynamic properties of the residual generators may be preferred, and therefore a method for finding sequential residual generators has been developed that accounts for these properties of the residual generator candidates. It is shown that only a small fraction of all residual generator candidates fulfill fundamental requirements, and thereby proves the value of systematic methods. Further, methods are devised for utilization of the residual generators, such as initialization of dynamic residual generators. A proposed method, considering the fault excitation in the residuals using the internal form of the residuals, significantly increases the diagnosis performance. A hybrid electric vehicle is used in a simulation study for demonstration, but the methods used are general in character and provides a basis when designing diagnosis systems for other complex systems.

1 INTRODUCTION

Design of a diagnosis system for a non-trivial real application is a complex engineering task involving many aspects (Blanke et al., 2006; Gentil et al., 2004; Cordier et al., 2004). Physical and semi-physical modeling will typically lead to a set of equations where some are algebraic and some are differential equations. Based on such a model of differential-algebraic equations (DAE), it is typically possible to construct many residuals to be used for fault detection and isolation (Blanke et al., 2006). How to select which of these residual generators to use is a non-trivial task, and further, before a diagnostic decision is made the residuals are often post-processed to form test quantities, e.g., by using the well-known CUSUM algorithm (Page, 1954).

When hybridizing a vehicle (Guzzella and Sciarretta, 2007), new components are added compared to a conventional vehicle, e.g. electric machines, battery, and power electronics. It is important to monitor these components due to safety issues and to avoid damaging components, especially the battery that is sensitive and costly. From the viewpoint of needed methodology for diagnosis design, it is fundamental that these systems by their very nature switch between combustion engine and electrical motor, and under these two main modes there are many sub-modes. This may typically result in that different residuals are switched on and off, resulting in design questions on e.g. reinitialization when being turned on.

The contributions in the paper concerns the selection of the residual generators to be used in the diagnosis system, and methods for utilizing these residual generators to make diagnostic decisions. The method used for generating residuals is based on computation sequences of the unknown variables (Staroswiecki and Declerck, 1989), and is called sequential residual generators by Svärd and Nyberg (2010). To this method, one would like to add engineering insights where intuition regarding noise properties may make a dynamic computation sequence leading to integration preferable compared to a sequence resulting in differentiation. Different computation sequences leads to different results regarding algebraic loops and uniqueness, which is exploited in the diagnosis design. Furthermore, in non-linear systems the fault excitation often depends on the operating mode of the system to be monitored, and a method for exploiting this when designing the test quantities is proposed. The method, that is one of the main contributions in the paper, is based on the internal form of the residual generators and is exemplified for the CUSUM algorithm. Finally, all these aspects, i.e. the investigation of all sequential residual generator candidates and utilization of a set of these, are illustrated on an industrial example in a simulation study for a hybrid electric vehicle.

2 RESIDUAL GENERATOR CONSTRUCTION

The objective of this section is to introduce, and make slight modifications to, the model based residual generation technique that is analyzed further in the fol-

lowing sections. Examples of proposed techniques to generate residuals are parity equations (Chow and Willsky, 1984), variable elimination (Staroswiecki and Comtet-Varga, 2001), parameter estimation (Isermann, 2006), and state-observer (Frank, 1994) techniques. As stated above, sequential residual generators are used here and the basic principle is straightforward (Staroswiecki and Declerck, 1989; Svärd and Nyberg, 2010). A sequential residual generator consists of a set of equations with analytical redundancy where a subset of the equations are used to solve for the unknown variables included in these equations, and then the redundant equations are used to check consistency between the observations and the set of model equations. The basic principle is illustrated with a small example.

Example 1. *Consider the model*

$$e_1 : x_1 = g_1(x_2)$$

$$e_2 : x_2 = g_2(u)$$

$$e_3 : y = x_1$$

with 3 equations, 2 unknown variables x_1 and x_2 , known input variable u , measurement variable y , and non-linear functions g_i . Here, both e_1 and e_3 can be chosen as the consistency relation, or analytical redundancy relation (ARR) (Staroswiecki and Comtet-Varga, 2001), resulting in the two residual generators

$$x_2 := g_2(u)$$

$$x_1 := y$$

$$r_1 := x_1 - g_1(x_2)$$

$$x_2 := g_2(u)$$

$$x_1 := g_1(x_2)$$

$$r_2 := y - x_1.$$

respectively.

In the simple example above, the different choice of consistency relation resulted in equivalent residual generators. However, in general the choice may strongly affect the performance of the diagnosis system, which is shown in Sections 4 and 6 using the hybrid vehicle as an example.

2.1 SEQUENTIAL RESIDUAL GENERATION BY STRUCTURAL ANALYSIS

This presentation will be based on the notation in Svärd and Nyberg (2010) and a brief outline of the approach is included here. The system model is denoted by $M(E, X, Z)$ where $E = \{e_1, e_2, \dots, e_m\}$ is the set of equations, $X = \{x_1, x_2, \dots, x_n\}$ the set of unknown variables, and $Z = \{z_1, z_2, \dots, z_r\}$ the set of known variables. It is assumed that the set of equations is expressed as a semi-explicit DAE, i.e.

$$\dot{x}_d = f(x_d, x_a, z) \tag{1a}$$

$$0 = g(x_d, x_a, z) \tag{1b}$$

where z is a vector of known variables and x_d are the dynamic variables and x_a the algebraic variables, both vectors of unknown variables x_i . The semi-explicit form leads to that no more than one differentiated variable is included in each equation.

In accordance with the example above, a sequential residual generator is based on a set of equations $\bar{E} \subseteq E$ with analytical redundancy where the set of unknown variables in \bar{E} is denoted by $X' \subseteq X$. In this presentation, subset minimal sets of equations with structural redundancy 1 will be considered, i.e., $|\bar{E}| = |X'| + 1$ and no proper subset $\bar{E} \subsetneq \bar{E}$ has structural redundancy > 0 . A sequential residual generator based on $M(\bar{E}, X', Z)$ then consists of a consistency relation $e_i \in \bar{E}$ and the just-determined set of equations $E' = \bar{E} \setminus \{e_i\}$ that is used to solve for the unknown variables X' . The key steps in the residual generator design is the choice of consistency relation and the solving procedure for the just-determined set of equations.

Structural methods analyze the model structure, i.e., only information about which variables that are included in each equation. This is typically represented by the incidence matrix of a bi-partite graph with node sets E and $X \cup Z$, and each edge in the graph corresponds to a variable dependency in an equation. As an example, the structure of the model in Example 1 can be represented by the incidence matrix

	x_1	x_2	y	u
e_1	X	X		
e_2		X		X
e_3	X		X	

For the first step in finding a sequential residual generator, efficient methods from structural analysis (Krysander et al., 2008) exists to, based on the model structure, find all minimal sets of equations with redundancy 1. Such sets are referred to as a *Minimally Structurally Overdetermined* (MSO) set of equations, i.e., a set of equations with 1 more equation than variables and where no proper subset of the equations is structurally overdetermined.

In the second step, when a consistency relation has been selected, the square non-linear system of equations E' need to be solved. Also here tools from structural analysis is useful. A *matching* in a bi-partite graph is a subset of edges with no common nodes, i.e., a pairing of equations and variables. From a maximal matching, a *computation sequence* for the unknown variables X' can be obtained.

Definition 1 (Computation sequence). *A computation sequence for $M(E', X', Z)$ is an ordered set $C = \{(V_1, E_1), \dots, (V_n, E_n)\}$ where $V_i \subseteq X'$ and $E_i \subseteq E'$. The order to compute the unknown variables, X' , in the set of equations, E' , is defined by the order of appearance in the set.*

For example, the computation sequence for the second case in Example 1 is

$$C = \{(\{x_2\}, \{e_2\}), (\{x_1\}, \{e_1\})\} \quad (2)$$

indicating that x_2 is computed from e_2 and x_1 is then computed from e_1 . In case of computational loops, i.e., cases where several variables has to be solved for concurrently, sets E_i and V_i are nonsingletons.

2.2 AN ALGORITHM

The methodology of investigating all residual generator candidates is used in this paper, and the function `FINDRESIDUALGENERATORS` that finds and investigates the properties of these in Svärd and Nyberg (2010) is recalled below. The input variables to the function are the model equations, E , the unknown variables, X , and an algebraic equation solver, \mathcal{T} . \mathcal{T} solves an equation system, and $\mathcal{T}(\mathcal{C})$ gives the expressions to be used to compute the unknown variables. For example, consider \mathcal{C} in (2), $\mathcal{T}(\mathcal{C}) = \{x_2 := g_2(u), x_1 := g_1(x_2)\}$. A sequential residual generator $r_j \in R$ consists of a computation sequence and a consistency relation, i.e., $r_j = (\mathcal{T}(\mathcal{C}), e_i)$. The call $\text{var}_X(\bar{E})$ returns the variables included in \bar{E} , the function `FINDALLMSOS` finds all MSOs given E and X , and `FINDCOMPUTATIONSEQUENCE` investigates the properties of the computation sequence of the unknown variables (see Svärd and Nyberg (2010) for details). This is done by (1) using the structural model to find a computation sequence, (2) investigate how dynamic equations are used in the computation sequence (see Section 2.3 for details), and (3) investigate if the algebraic solver \mathcal{T} is able to find expressions for the unknowns variables. If there is no realizable algebraic solution the computation sequence, \mathcal{C} , is empty.

```

1: function FINDRESIDUALGENERATORS( $E, X, \mathcal{T}$ )
2:    $R := \emptyset$ ;
3:    $MSOs := \text{FINDALLMSOS}(E, X)$ ;
4:   for all  $\bar{E} \in MSOs$  do
5:      $X' := \text{var}_X(\bar{E})$ ;
6:     for all  $e_i \in \bar{E}$  do
7:        $E' := \bar{E} \setminus e_i$ 
8:        $\mathcal{C} := \text{FINDCOMPUTATIONSEQUENCE}(E', X', \mathcal{T})$ 
9:       if  $\mathcal{C} \neq \emptyset$  then
10:         $R := R \cup \{(\mathcal{T}(\mathcal{C}), e_i)\}$ 
11:       end if
12:     end for
13:   end for
14:   return  $R$ 
15: end function

```

2.3 DYNAMIC MODELS

The model used in this investigation is based on equations in the form (1). To make analysis of dynamic models explicit, relations between a variable $x_i \in X$

and \dot{x}_i is included in the model using an equation

$$\frac{d}{dt}x_i = \dot{x}_i \quad (3)$$

Such an equation can, in a particular computation sequence, be solved in either of two directions (Blanke et al., 2006; Frisk et al., 2012)

- **derivative causality** is when x_i in (3) is differentiated to obtain \dot{x}_i , i.e., $\dot{x}_i := \frac{d}{dt}x_i$.
- **integral causality** is when \dot{x}_i in (3) is integrated to obtain x_i , i.e. $x_i := \int \dot{x}_i dt + x_0$, where x_0 is the initial value of x_i .

A computation sequence is said to be in derivative causality if all dynamic constraints are solved in derivative causality and similar for integral causality. If both derivative and integral causality is used in the computations, the sequence is said to be in mixed causality.

2.4 MODIFICATION TO THE ALGORITHM TO HANDLE DYNAMIC CONSISTENCY RELATIONS

The algorithm used to find the properties of the residual generator candidates is described in Section 2.2 and (Svärd and Nyberg, 2010). To also consider dynamics in the consistency relation, a modification is made to the original algorithm.

An illustrative example is given below to exemplify the case when there is dynamics in the set of equations to be used to construct the residual generator.

$$e_1 : \dot{x}_1 - u = 0 \quad (4a)$$

$$e_2 : x_1 - y = 0 \quad (4b)$$

If e_2 is selected to be the consistency relation, i.e. $e_i = e_2$, then e_1 is used to compute x_1 , and integral causality is used since \dot{x}_1 is included in e_1 . If, on the other hand, e_1 is selected as the consistency relation, i.e. $e_i = e_1$, then e_2 is used to find x_1 , and the algorithm given in Section 2.2 gives that neither integral or derivative causality is used since there is no dynamic equation included in the computation sequence (i.e. e_2 in this case). However, to be able to use e_1 as the consistency relation, x_1 has to be differentiated, and to also consider this case and achieve the causalities given in Table 1, a slightly modified algorithm is now presented below. There, the call $\text{var}_D(e_i)$ returns the differentiated variable, if there is any, in equation e_i . For example, the call $\text{var}_D(e_1)$, where e_1 is given in (4a), results in \dot{x}_1 . Due to that the model is given in semi-explicit form, a differentiated variable is only included once in the set of equations. This implies that if a differentiated variable is included in the consistency relation, this variable is only known in its undifferentiated form from the computation sequence. Therefore derivative causality is used in such a case. The function

Table 1: The causalities returned from the algorithms described for the example presented in (4).

Consistency relation	Original	Modified
e_1	Static	Derivative
e_2	Integral	Integral

ISDERIVATIVEPERMITTED(\mathcal{D}) is true if derivative or mixed causality is to be used, otherwise false. The parameter \mathcal{D} is an input parameter and includes information about what causality that is permitted to solve dynamic equations.

In FINDRESIDUALGENERATORS all realizable residual generators are saved in R , but in the modified algorithm only the realizable residual generators fulfilling the constraint regarding the causality is to be stored in R . If \mathcal{C} fulfills the requirements in \mathcal{D} is investigated in the function ISCAUSALITYOK(\mathcal{C}, \mathcal{D}).

```

1: function FINDRESIDUALGENERATORSMOD( $E, X, \mathcal{T}, \mathcal{D}$ )
2:    $R := \emptyset$ ;
3:    $MSSOs := \text{FINDALLMSOs}(E, X)$ ;
4:   for all  $\bar{E} \in MSSOs$  do
5:      $X' := \text{var}_X(\bar{E})$ ;
6:     for all  $e_i \in \bar{E}$  do
7:       if ISDERIVATIVEPERMITTED( $\mathcal{D}$ ) or  $\text{var}_{\mathcal{D}}(e_i) = \emptyset$  then
8:          $E' := \bar{E} \setminus e_i$ 
9:          $\mathcal{C} := \text{FINDCOMPUTATIONSEQUENCE}(E', X', \mathcal{T})$ 
10:        if  $\mathcal{C} \neq \emptyset$  and ISCAUSALITYOK( $\mathcal{C}, \mathcal{D}$ ) then
11:           $R := R \cup \{(\mathcal{T}(\mathcal{C}), e_i)\}$ 
12:        end if
13:      end if
14:    end for
15:  end for
16:  return  $R$ 
17: end function

```

3 VEHICLE MODEL

A model of a hybrid electric vehicle (HEV), 32 equations, is used together with sensor models, five equations, and fault models, five equations. The complete model consisting of the truck model, sensor models and faults are described in Appendix A, Sections 3.2 and 3.3, respectively. For the sake of the analysis to come, relevant parts of the vehicle model in Appendix A are pointed at in Section 3.1.

3.1 POWERTRAIN MODEL

The modeled vehicle is a long haulage electric parallel hybrid truck, with the connection of the electrical and conventional parts of the powertrain located

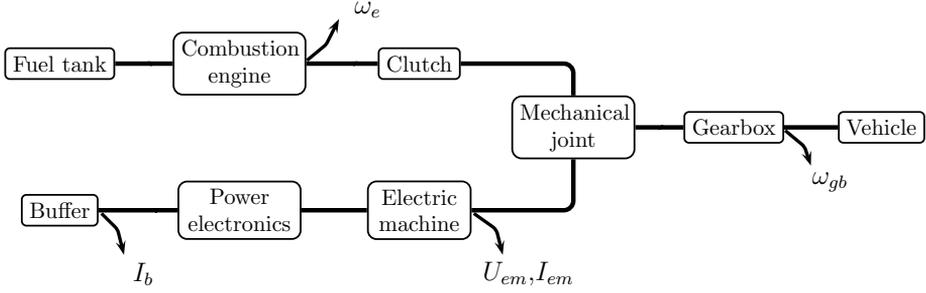


Figure 1: The configuration of the modeled vehicle. The arrows and signals indicate the sensors used in the diagnosis systems described in Sections 4.2 and 6.

between the clutch and the gearbox, see Figure 1. The models of the components in the vehicle are based on the model library developed in the MATLAB/SIMULINK based Center for Automotive Propulsion Simulation (Fredriksson et al., 2006) and the QSS toolbox (Guzzella and Amstutz, 1999). Modifications to these models are carried out to model a truck instead of a passenger car, and to include the possibility to induce faults in the models as well as to add sensor noise. The energy management is based on a heuristic control strategy (see Sundström (2011); Sundström et al. (2010) for details), and a Willans model is used for the combustion engine (Guzzella and Sciarretta, 2007; Rizzoni et al., 1999). All model equations for the powertrain are given in Appendix A and are described in Sundström (2011).

BATTERY

The battery is modeled using a Thévenin equivalence circuit, i.e. a voltage source, U_{oc} , and an inner resistance, R_b , connected in series (Reddy, 2011). The battery voltage on the terminals for a current I_b is

$$U_b = U_{oc}(SoC) - R_b I_b \quad (5)$$

and the state of charge, SoC , is based on I_b and the capacity of the battery, Q_b

$$SoC = SoC_0 - \frac{1}{Q_b} \int I_b dt, \quad SoC \in [0, 1] \quad (6)$$

where SoC_0 is the initial state of charge of the battery.

The modeled battery pack is based on 200 cells. There are two chains of cells connected in parallel, and the cells used are denoted as the “10-Mile PHEV” in Nelson et al. (2007). The resistance for the battery pack is assumed to be a constant based on the tabulated resistance at 50% SoC ($R_b = 0.023\Omega$). The total capacity of the battery pack is 33.2 Ah, the $U_{oc}(SoC = 0.5)$ is 250 V, leading to an energy storage capacity of approximately 8 kWh. The weight of the battery is 150 kg, of this the weight of the cells is 129 kg.

POWER ELECTRONICS

The model of the power electronics is assumed to be an ideal component

$$P_b = P_{em} \Leftrightarrow I_b U_b = I_{em} U_{em} \quad (7)$$

where P_b and P_{em} are the electrical powers from the battery and motor, and U_b , I_b , U_{em} , and I_{em} the battery and electric machine voltages and currents.

ELECTRIC MACHINE

The electric machine is able to convert electric power to mechanical power and vice versa. A voltage, U_{em} , is applied on the component, resulting in a torque on the outgoing shaft. The torque, T_{em} , is proportional to the armature current, I_{em} , with the torque constant k_a , except for the frictional losses that are assumed to be proportional to the speed of the machine, ω_{em} , with the gain $c_{em,f}$. The speed constant k_i is used to model the electromotive force

$$I_{em} = \frac{U_{em} - k_i \omega_{em}}{R_{em}} \quad (8a)$$

$$T_{em} = I_{em} k_a - c_{em,f} \omega_{em} \quad (8b)$$

where R_{em} is the resistance.

In an ideal Permanent Magnet Synchronous Machine (PMSM), that is a common machine type in HEVs due to its high efficiency (Chau et al., 2008), k_i and k_a are equal. These constants are defined by $K\phi$, where K is a machine constant that depends on design parameters of the machine, and ϕ is the magnetic flux produced by the stator. One way to model the losses of the machine is to use $k_a \neq k_i$ (Guzzella and Sciarretta, 2007). This is here done by

$$k_a = k_i \eta_{em,0}^{\text{sign}\{I_{em}\}} \quad (8c)$$

where $\eta_{em,0}$ is the efficiency and $\text{sign}\{\cdot\}$ the signum function.

The model is parametrized as a 59 kW DC machine with constant magnetic flux.

CLUTCH

There is a model of the clutch to handle take off and gear shifts. The clutch model is modified compared to the model presented in e.g. Eriksson (2001), to decrease the stiffness of the simulation model when there is slip in the clutch. The delivered torque from the clutch is given by

$$T_c = u_c T_e \quad (9)$$

where u_c is the position of the clutch and T_e is the delivered torque from the engine. The engine speed is assumed to be

$$\omega_e = \begin{cases} \omega_{e,idle}, & \omega_{mj} < \omega_{e,idle} \\ \omega_{mj}, & \omega_{mj} \geq \omega_{e,idle} \end{cases} \quad (10)$$

where ω_e , $\omega_{e,idle}$, and ω_{mj} are the engine speed, engine idle speed, and speed in the mechanical joint, respectively.

VEHICLE

In the vehicle, that has a mass m_v and a lumped inertia J_{tot} , the output shaft from the gearbox is connected to the final gear and finally to the wheels. The gear ratio in the final gear is u_f , and the angular acceleration of the wheels is calculated based on the resulting torque, T_{net} , acting on the wheels, whose radius is r_w

$$T_{net} = T_{gb}u_f - T_d - T_r - T_b - T_g \quad (11a)$$

$$\dot{\omega}_w = \frac{T_{net}}{J_{tot}u_f^2 + m_v r_w^2} \quad (11b)$$

where T_{gb} , T_d , T_r , T_b , and T_g are the torques from the gearbox, air drag, rolling resistance, mechanical brakes, and slope of the road, respectively.

3.2 SENSORS

The five sensors available for the diagnosis system used in the analysis of the model in Section 4.2 and the simulation study described in Section 6 are shown in Figure 1. The engine speed is ω_e , and ω_{gb} is the speed of the outgoing shaft from the gearbox. The naming convention of a sensor signal is the name of the measured variable with the extension SENS in the subscript, e.g. $\omega_{gb,sens}$.

The measurement signal, y , is given by the noise free signal, y^* , added with noise, $\tilde{\nu}$

$$y = y^* + \tilde{\nu} \quad (12)$$

3.3 INDUCED FAULTS

To model that the battery, power electronics, and the electric machine may break down, two parameter values and two voltages in these models have the possibility to be modified. Note that these faults are only examples of how a fault in these components can be represented in the model. The following modifications of the signals are introduced to model the faults where the nominal signals are denoted by the superscript *nom*:

$$f_{b,sc} : U_b = (1 + f_{b,sc}) U_b^{nom} \quad (13a)$$

$$f_{pe} : U_{em} = (1 + f_{pe}) U_{em}^{nom} \quad (13b)$$

$$f_{em,\eta} : \eta_{em,0} = (1 + f_{em,\eta}) \eta_{em,0}^{nom} \quad (13c)$$

$$f_{em,R} : R_{em} = (1 + f_{em,R}) R_{em}^{nom} \quad (13d)$$

where $f_{b,sc}$ models that all cells in the battery are not used due to an internal short circuit, f_{pe} is a fault in the power electronics, and $f_{em,\eta}$ and $f_{em,R}$ are two faults in the electric machine.

Table 2: Values of the faults induced in the model.

Faults	Value
$f_{em,\eta}$	-0.5
$f_{em,R}$	-0.5
f_{pe}	-0.5
$f_{b,sc}$	-0.5
$f_{em,U,sens}$	20 V

In addition to the four faults modes described in (13a)-(13d), the voltage sensor in the electric machine is modeled to possibly include an offset fault, $f_{em,U,sens}$

$$U_{em,sens} = U_{em} + f_{em,U,sens} \quad (13e)$$

When a fault is induced in the model, the value of the fault is given in Table 2.

4 SELECTION OF CONSISTENCY RELATION

Each equation in an overdetermined set of equations can be selected as the consistency relation used to construct a sequential residual generator. However, as described above, the algebraic and dynamic properties of the residual generator may vary with the choice of consistency relation, even though they are based on the same set of model equations. These properties will now be studied. The HEV model in Section 3 is used, but general conclusions will be made. The presentation is structured so that, first, a simplified but realistic example is used to illustrate the possibility to avoid algebraic loops by consistency relation selection. Secondly, all residual generator candidates based on all MSOs for the vehicle are analyzed regarding algebraic loops and uniqueness, as well as how the causality of the dynamic equations are used. In the latter case, a quantitative analysis shows that the number of possible residual generators differs significantly, thus demonstrating the advantages of the systematic approach.

4.1 AVOIDING ALGEBRAIC LOOPS BY CONSISTENCY RELATION SELECTION

Algebraic loops, as well as multiple solutions of the unknown variables, may occur in computation sequences (Blanke et al., 2006; Katsillis and Chantler, 1997). There are several numerical and analytical methods available to solve algebraic loops. Linear loops are e.g. easily solved, but non-linear algebraic loops may demand a large computation effort to solve and a solution is not always obtained. Thus, it is sensible to avoid algebraic loops if possible and here the computation sequences are analyzed with respect to algebraic loops.

The possibility to find a unique residual generator given a set of equations in an MSO varies with the chosen consistency relation. To illustrate this an MSO based on the vehicle model presented in Section 3.1 is used, and since the MSO consists of many equations three simplifications are made in this example:

- the torques due to losses in the vehicle are lumped and denoted T_l
- the torques T_e and T_l are assumed to be known
- the angular speeds in the powertrain are equal and are denoted ω

resulting in the following set of equations

$$\begin{aligned}
 e_1 : \quad & T_e + \underbrace{\eta_{em,0}^{\text{sign}\{I_{em}\}} k_i I_{em} - c_{em,f} \omega - T_l(\omega)}_{T_{em}} - J_{tot} \dot{\omega} = 0 \\
 e_2 : \quad & \frac{U_{em} - \omega k_i}{R_{em}} - I_{em} = 0 \\
 e_3 : \quad & I_b U_b - I_{em} U_{em} = 0 \\
 e_4 : \quad & U_{oc} - U_b - R_b I_b = 0 \\
 e_5 : \quad & \frac{d}{dt} \omega - \dot{\omega} = 0 \\
 e_6 : \quad & U_b - y_1 = 0 \\
 e_7 : \quad & \omega - y_2 = 0
 \end{aligned} \tag{14}$$

where y_1 and y_2 are sensor signals, and ω , $\dot{\omega}$, U_b , I_b , U_{em} , and I_{em} are the six unknown variables.

If e_1 in (14) is selected as the consistency relation, the permuted structural model of the just-determined part, i.e. $\{e_2 - e_7\}$, is given in Table 3. The corresponding computation sequence is

$$\mathcal{C} = \{(\{\omega\}, \{e_7\}), (\{U_b\}, \{e_6\}), (\{\dot{\omega}\}, \{e_5\}), (\{I_b\}, \{e_4\}), (\{I_{em}, U_{em}\}, \{e_2, e_3\})\} \tag{15}$$

The pair $(\{I_{em}, U_{em}\}, \{e_2, e_3\})$ indicates that there is an algebraic loop, that also can be seen in Table 3. This loop has the non-unique solution

$$I_{em} = -\frac{\omega k_i}{2R_{em}} \pm \sqrt{\left(\frac{\omega k_i}{2R_{em}}\right)^2 + \frac{I_b U_b}{R_{em}}} \tag{16}$$

If one of e_2 , e_3 , or e_4 is used as consistency relation instead of e_1 , there is no algebraic loop in the just-determined part. The key to avoid the algebraic loop in this example is to find a computation sequence where I_{em} is computed in e_1 . The computation sequence of the unknown variables if e.g. e_2 is used as the consistency relation will be

$$\mathcal{C} = \{(\{\omega\}, \{e_7\}), (\{U_b\}, \{e_6\}), (\{\dot{\omega}\}, \{e_5\}), (\{I_b\}, \{e_4\}), (\{I_{em}\}, \{e_1\}), (\{U_{em}\}, \{e_3\})\} \tag{17}$$

Table 3: Permuted structural model of the system given in (14) except e_1 that is chosen to the consistency relation. Equations e_2 and e_3 form an algebraic loop for I_{em} and U_{em} .

	ω	U_b	$\dot{\omega}$	I_b	U_{em}	I_{em}
e_7	X					
e_6		X				
e_5	X		X			
e_4		X		X		
e_3		X		X	X	X
e_2	X				X	X

Since the equations in the substitution chain are uniquely solvable, a unique residual generator is found.

Remark 1. *Note that since consistency based diagnosis is used, it is possible to construct a test that is based on a residual generator with several solutions. As long as at least one of the possible residuals, r_1, r_2, \dots, r_n where n is the number of solutions, is close to zero, the residual will not react on the fault*

$$|r(t_k)| = \min\{|r_1(t_k)|, |r_2(t_k)|, \dots, |r_n(t_k)|\}, n \geq 2$$

However, the computation complexity of the system increases if more than one residual are to be evaluated in a diagnostic test.

Remark 2. *Sensor equations are often selected to be the consistency relation, but with respect to algebraic loops it is in general preferable to include such an equation in the computation sequence. The reason is that a sensor equation only includes one unknown variable, and therefore can never be part of a loop in the computation sequence. This is exemplified above, where algebraic loops occur if any sensor equation is selected as consistency relation.*

SERIES WOUND ELECTRIC MACHINE

In the example above the magnetic field, ϕ , created by the stator or armature in the electric machine is assumed to be constant. This is the case in permanent magnet synchronous machines, which is the machine type mainly used for vehicle propulsion in HEVs. But if a series wound machine is used instead, that e.g. is used in starter motors (Hambley, 2005), then e_1 and e_2 in (14) are modified to

$$\begin{aligned}
 e_1 : \quad T_e + \underbrace{\eta_{em,0}^{\text{sign}\{I_{em}\}} k_i I_{em}^2 - c_{em,f} \omega - T_l(\omega)}_{T_{em}} - J_{tot} \dot{\omega} &= 0 \\
 e_2 : \quad \frac{U_{em}}{R_{em} + \omega k_i} - I_{em} &= 0
 \end{aligned} \tag{18}$$

due to that ϕ increases linearly with the current in the rotor and stator according to

$$\phi = k_a I_{em} = \eta_{em}^{\text{sign}\{I_{em}\}} k_i I_{em} \quad (19)$$

The same variables are included in each equation in (14) and (18), and therefore the structural models are the same for the two machine types. It is however not possible to chose a consistency relation that results in a unique expression for the residual generator for the latter machine type. There are three equations, e_1 , e_2 , and e_3 , that can be used to find a matching for I_{em} . If e_2 or e_3 is used to compute I_{em} there is a non-linear algebraic loop that has multiple solutions (similar to (16)), and I_{em} is quadratic in e_1 .

4.2 PROPERTIES OF THE SEQUENTIAL RESIDUAL GENERATORS CANDIDATES

Now return to the task of finding candidates for sequential residual generators. Given the powertrain model presented in Section 3.1 and the five sensors available when designing the diagnosis systems (see Figure 1), 99 MSOs are found. For these MSOs there are 2667 sequential residuals generator candidates that are investigated using the algorithm described in Section 2.4. The number of these residual generator candidates that fulfill different causality requirements, i.e. mixed, integral, derivative, and static, are presented in Figure 2. In the figure the white bars indicate the number of model equations, i.e. the number of residual generator candidates, included in each MSO. All equations are not invertible and the number of residual generators that are realizable and using corresponding causality are given by the gray bars. The black bars represent how many of the equations that can be selected as a consistency relation to achieve a unique residual generator for each causality. Note that static residual generators are included in Figures 2a-2c, and that pure derivative or integral causality residual generators are included in Figure 2a.

For MSOs 15-99 there is only a small fraction of the equations that can be used as a consistency relation if only derivative or integral causality is used, and there are five MSOs that it is not possible to construct a unique residual generator by using mixed or integral causality (see the non-black bars in Figure 2a and 2b). However, if the constraint regarding unique residual generators is relaxed to allow multiple solutions, as for example in (16), it is possible to design tests based on all MSOs if integral or mixed causality is used (all MSOs in Figures 2a and 2b contain black or gray bars).

When derivative causality is used to solve the dynamic equations, it is found that only 43 MSOs can be used to construct diagnosis tests (the black bars in Figure 2c). Additional investigations show that this leads to that full fault isolability is not achieved. However, full fault isolability is structurally achieved when mixed or integral causality is used.

Table 4 includes the same type of results as Figure 2, but instead gives aggregated numbers. The numbers in parenthesis is the result, when adequate,

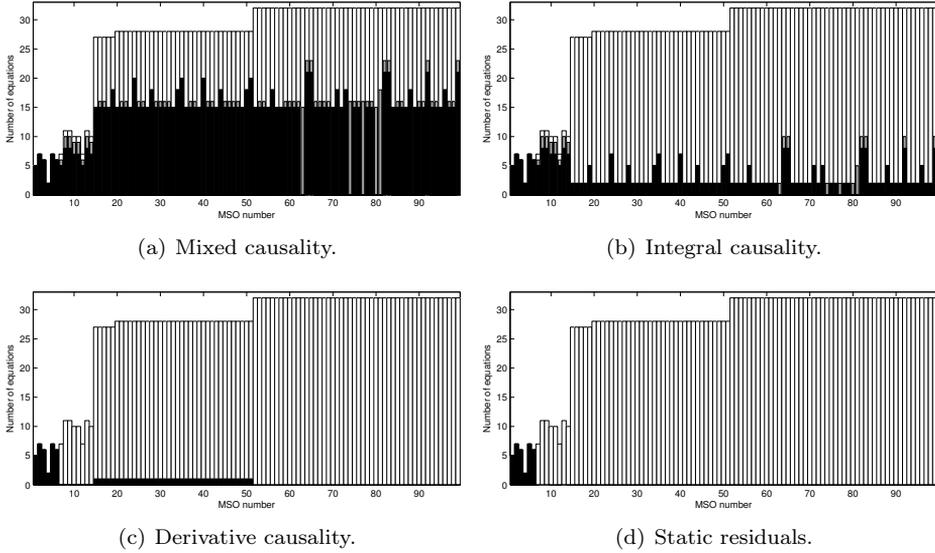


Figure 2: The white bars indicate the number of equations in each MSO in the diagnosis systems, the black bars the number of equations that can be selected to consistency relations to fulfill the causality constraint and that a unique sequential residual generator is to be found, and the gray bars if the constraint regarding unique residual generators is relaxed to that a residual generator is to be found (i.e. unique or multiple solutions of the unknown variables).

from the original algorithm `FINDRESIDUALGENERATORS` described in Section 2.2 to be compared with the results from `FINDRESIDUALGENERATORSMOD` described in Section 2.4, and it is clear that including the consistency relation in the analysis affects the results. I, D, and S indicate that integral causality, derivative causality, and static expressions are used.

4.3 SUMMARY AND DISCUSSION

In the illustrative example in Section 4.1 it was shown that the occurrence of algebraic loops varies with the selection of consistency relation. It was also shown that the possibility to find a unique residual generator depends on the model the residual generator is based on, where in this specific case it is possible to find a unique residual generator using a permanent magnet electric machine, but not if a series wound electric machine is used.

For selection of sequential residual generators, the main result so far is Figure 2 and Table 4. The analysis of the entire vehicle model gives that there is a small fraction of the residual generators that uses integral and derivative causality (compare the black and white bars in Figures 2b and 2c). In general

Table 4: The number of MSOs and sequential residual generators with different algebraic and dynamic properties. In total there are 2667 residual generator candidates. Combinations of S, D, and I occur, and e.g. SI indicates that both static expressions and integral causality may be used. The figures in parenthesis is the result, when adequate, from the original algorithm described in Section 2.2.

Algebraic constraint	Causality	# MSO sets	# res. gen.
Unique residuals	SDI	94	1374
	SI	94 (94)	334 (414)
	SD	43	70
	S	6 (6)	33 (33)
Realizable residuals	SDI	99	1520
	SI	99 (99)	373 (458)
	SD	43	70
	S	6 (6)	33 (33)

there are few choices of consistency relations when designing a sequential residual from an MSO to achieve certain properties regarding algebraic solutions and causality. Therefore it can be stated that systematic methods are valuable to investigate the properties of the sequential residual generator candidates.

5 METHODS FOR UTILIZATION OF RESIDUAL GENERATORS AND TEST QUANTITIES

There are several important issues when using the residual generators and test quantities in a diagnosis system, and in this section three particular topics are discussed. First a technique is presented to avoid differentiation in the consistency relation, secondly reinitialization of dynamic residual generators, and finally a way to increase test performance by taking fault sensitivity into consideration. The residual generators are constructed using the algorithm given in Section 2.4, and a general model given as a semi-explicit DAE is considered.

5.1 AVOID DIFFERENTIATING IN THE CONSISTENCY RELATION USING A STATE TRANSFORMATION

There are several ways of numerically differentiating a signal, see e.g. Barford et al. (1999); Frisk and Åslund (2005). However, here a method is described that leads to that no differentiated variable needs to be computed in the redundant equation.

Since the model is given as a semi-explicit DAE, see (1), the differentiated variable is included linearly in the consistency relation. This leads to a residual generator that can be expressed in the general form

$$\tilde{r} = \dot{x}_i + b(z) \quad (20a)$$

where x_i is the variable that occurs in the consistency relation in its differentiated form and $b(z)$ an arbitrary function of known signals. Residuals including differentiated variables are often filtered, and by filtering \tilde{r} in (20a) the residual r is obtained

$$r = \frac{\alpha}{p + \alpha} \tilde{r} = \frac{\alpha}{p + \alpha} (\dot{x}_i + b(z)) \tag{20b}$$

where p is the differentiation operator and $\alpha > 0$ for stability.

Now, it is possible to compute r in (20b), without calculating a differentiated signal using a transformation and state the residual generator in state-space form. The conditions for this to be possible is that b a function of known variables, and the residual is filtered as in (20b) (Frisk and Nyberg, 2001). Using the state

$$\Gamma = r - \alpha x_i \tag{20c}$$

we obtain that the residual generator in (20b) can be expressed as

$$\dot{\Gamma} = -\alpha\Gamma - \alpha^2 x_i + \alpha b \tag{20d}$$

$$r = \Gamma + \alpha x_i \tag{20e}$$

which is a computational form that does not include a differentiation of variable x_i . By using this methodology it is possible to always avoid differentiating a variable in a sequential residual generator if a signal only is to be differentiated in the consistency relation.

Drift in the state, that is a common issue when using integral causality, is not an issue using this methodology since the filter is asymptotically stable. This is the reason that this residual generator is not considered as using integral causality, even though a signal is integrated in the residual generator.

Remark 3. *Another possibility to avoid differentiating a variable to be used in the consistency relation, is by integrating the consistency relation. The residual instead of (20a) would e.g. be*

$$\bar{r}(t) = \int_{t_0}^t \tilde{r}(\tau) d\tau + r(t_0) = \alpha x(t) + \int_{t_0}^t b(z(\tau)) d\tau + b(z(t_0)) \tag{21}$$

where $b(z(t_0))$ is the initial value of $b(z)$. In this case the residual generator has the properties of integral causality, i.e. the initial value of the state has to be available and drift may occur in the integrators due to modeling errors.

Note that it is only differentiated variables in the consistency relation, and not in other parts of the MSO, that always can be avoided to compute by integration. To illustrate, consider the example MSO

$$e_1 : x_2 - u_1 = 0 \tag{22a}$$

$$e_2 : \dot{x}_2 - x_1 = 0 \tag{22b}$$

$$e_3 : \sin(x_1) - y_1 = 0 \tag{22c}$$

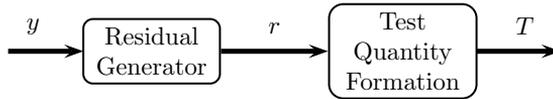


Figure 3: The residuals are computed from residual generators based on known signals y . The residuals are post processed to form test quantities, T , that are compared with thresholds to state if the tests have reacted.

Assume that integral causality is to be used in the residual generator, and that e_3 is used as the consistency relation. It is not possible to find a sequential residual generator that fulfills this requirement, since $\int x_1 dt$ is computed from e_2 in order to not differentiate x_2 , but x_1 needs to be known in the consistency relation.

5.2 INITIALIZATION OF STATES

When a dynamic residual generator is started or re-started, the states $\omega \in \mathbb{R}^n$ needs to be initialized. It is assumed that the residual generator is given in state space form

$$\dot{\omega} = f(\omega, z) \quad (23a)$$

$$r = h(\omega, z) \quad (23b)$$

where z is a vector of known variables. When the states ω are initialized the residual and its derivatives should ideally be zero in the fault free case. This leads to that the initial states $\omega(t_0)$ should be chosen such that

$$r(t_0) = \dot{r}(t_0) = \dots = r^{(n-1)}(t_0) = 0 \quad (24)$$

where t_0 is the time the residual is started.

The state in the first order system given in (20d) and (20e) is initialized using (24) as

$$\Gamma(t_0) = r(t_0) - \alpha x_i(t_0) = -\alpha x_i(t_0) \quad (25)$$

5.3 CONSIDER FAULT EXCITATION WHEN COMPUTING TEST QUANTITIES

To increase robustness of decision, residuals are often post processed by computing to what is here referred to test quantities, T , see Figure 3. A well known algorithm is CUSUM (Page, 1954; Gustafsson, 2000), that in this specific case can be stated as

$$T(t_k) = \max \{0, T(t_{k-1}) + |r(t_k)| - \nu\} \quad (26)$$

where ν is a design parameter that corresponds to the noise and model uncertainty in the residuals. The test reacts when T is above a threshold J .

The sensitivity of the residuals often varies with the operating condition of the system to be monitored. One way to improve the performance of the diagnosis system is to consider this aspect when computing the test quantities. A small academic example is used to illustrate the basic principle, and in Section 6.4 the approach is applied to the HEV model described in Section 3. Consider the system

$$x = u \quad (27a)$$

$$y = (1 + f)x \quad (27b)$$

where u and y are known signals, x an unknown signal, and f a fault. One residual generator of (27) is

$$r = y - u \quad (28)$$

and the internal form, i.e. how the faults affect the residual, of the residual generator is

$$r = u + f \cdot u - u = f \cdot u \quad (29)$$

From the internal form it is obvious that the residual sensitivity on the fault depends on the value of u .

To achieve good performance in the diagnosis system it is beneficial to only update T , as in (26), when the residual is excited by the fault. Figure 4 exemplifies this based on (27)-(29) where

$$u(t) = 10 \sin(t) \quad (30)$$

and f is a constant. In the figure it can be seen that the test does not react on the fault if T is updated also when there is low excitation in the residual, i.e. when r is close to zero even though there is a fault in the system. However, if T only is updated when the magnitude of y , that is assumed to be a good approximation of u , is above a threshold the fault is detected.

The offset parameter ν is here set to a fixed value. Instead of not updating T when there is low excitation in the residual, there is a possibility to modify ν . The disadvantage of this is that e.g. when $x = 0$, $\nu > 0$ in order to avoid false alarm due to sensor noise, leading to that T is decreasing for this operating point.

The conclusion from this simple example is that it is advantageous to use fault models to find the internal form of the residual generator, and to use this internal form to design the strategy for not updating the test quantity when there is low excitation in the system.

6 ILLUSTRATIVE DESIGNS AND SIMULATION STUDY

Two diagnosis systems, one based on mixed causality and one on integral causality, of the HEV described in Section 3 are evaluated to investigate the impact of different choices in the design of a diagnosis system. The general methodologies described in Section 5, e.g. initialization of the states and the internal form of the residual generators, are utilized.

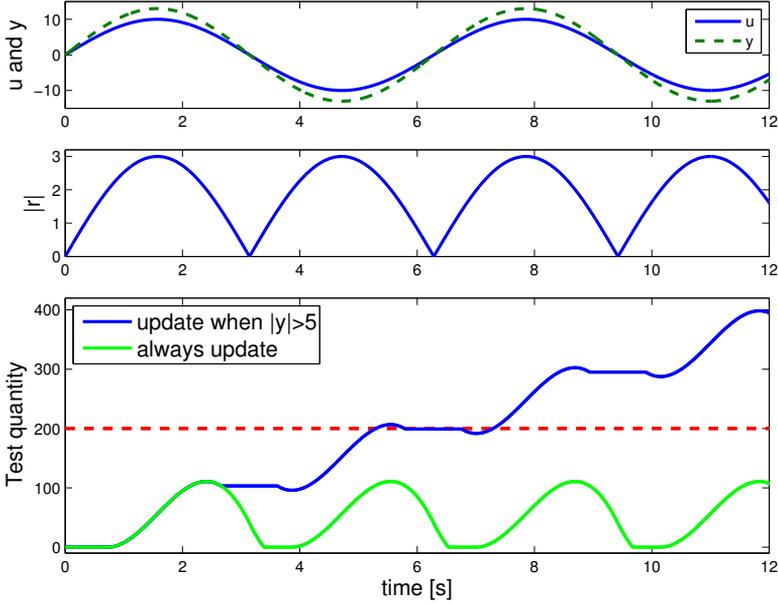


Figure 4: The variable $u(t)$ is a sinus function with amplitude 10 and $f=0.3$. For the selected ν and $J = 200$ the diagnostic test does not react on the fault if the test quantity, T , is updated all the time. However, if T only is updated when $|y| > 6$, the test reacts on the fault.

6.1 PROPERTIES OF DIAGNOSIS SYSTEMS USED IN SIMULATION STUDY

The two diagnosis systems use information from five sensors, see Figure 1, and are found using the algorithm `FINDRESIDUALGENERATORSMOD` from Section 2.4. If the original algorithm `FINDRESIDUALGENERATORS` were used, both these two systems would be classified to use integral causality, since the differentiation occurs in the consistency relation itself. Thereby these two systems illustrate the difference between the algorithms as discussed in Section 2.4.

MCDS

The diagnosis system based on mixed causality is denoted `MIXED CAUSALITY DIAGNOSIS SYSTEM`, or `MCDS` for short, and the equations in the computation sequences are uniquely solvable in the residual generators. The diagnosis system consists of four tests that achieves full structural single fault isolability (Krysander and Frisk, 2008) of the five faults. Each test is based on an MSO that is used to construct a sequential residual generator, that are given in Appendix B in

(47), (48), (51), and (52). Test 1 is static, Test 2 is based on integral causality, and Tests 3 and 4 are based on mixed causality. In the mixed causality tests the mass of consumed fuel, m_f , is solved by

$$m_f(t) = \int_{t_0}^t \dot{m}_f(\tau) d\tau + m_f(t_0) \quad (31)$$

Equation (11) is used as consistency relation, but the different torques are computed based on different sensors in the two tests. In the consistency relation w_{gb} is differentiated to compute $\dot{\omega}_{gb}$, resulting in that derivative causality is used.

For the computation of the residual, the reformulation in (20) is used and the residual is expressed in the form

$$\tilde{r} = \dot{\omega}_{gb} + b \quad (32a)$$

The residual generators are filtered and transformed according to (20b)-(20c), that results in

$$\dot{\Gamma} = -\alpha\Gamma - \alpha^2\omega_{gb} + \alpha b \quad (32b)$$

$$r = \Gamma + \alpha\omega_{gb} \quad (32c)$$

The algebraic loop for I_{em} and U_{em} considered in Section 4.1 is not an issue in these two residual generators, since U_{em} is known without using any of e_2 and e_3 in (14). The required voltage from the power electronics is known in Test 3, and the sensor measuring U_{em} is available in Test 4.

ICDS

A diagnosis system based on integral causality and unique expressions of the unknown variables in the computation sequences is designed, and this system is denoted INTEGRAL CAUSALITY DIAGNOSIS SYSTEM, or ICDS for short. It is possible to use the same sets of equations that are used in MCDS, and using the same MSOs the structural isolability properties are the same. Tests 1 and 2 in MCDS can also be used in ICDS, see (47) and (48) for corresponding residual generators, while different consistency relations are to be selected in Tests 3 and 4 in the ICDS since the consistency relations selected in MCDS result in mixed causality.

Tests 3 and 4 are based on MSOs with 29 and 32 equations respectively, but only two of these, e_{32} and e_{37} in Appendix A, are possible to select as consistency relations in this system

$$e_{32} : \omega_w = \frac{\omega_{gb}}{u_f} \quad (33)$$

$$e_{37} : \omega_{gb} = \omega_{gb,sens} \quad (34)$$

When using any of these as a consistency relation, $\dot{\omega}_w$ is calculated using (11), and ω_w by integrating this signal

$$\omega_w(t) = \int_{t_0}^t \dot{\omega}_w(\tau) d\tau + \omega_w(t_0) \quad (35)$$

In the residual generators used in both Tests 3 and 4, (33) is used as the consistency relation

$$r = \omega_{gb,sens} - \omega_w u_f \quad (36)$$

and the residual generators are presented in (49) and (50). Note that the only difference between these residual generators and the ones used for Tests 3 and 4 in MCDS are the last four equations in the computation sequences and the consistency relations. In the residual generators in MCDS the consistency relation is e_{30} and the last part of the computation sequence ($\{\omega_w\}, \{e_{32}\}$), ($\{T_d\}, \{e_{25}\}$), ($\{T_r\}, \{e_{26}\}$), ($\{T_{net}\}, \{e_{29}\}$), while e_{32} is used as the consistency relation in Tests 3 and 4 in ICDS and there is a dynamic loop in the computation sequence ($\{T_d, T_r, T_{net}, \omega_w\}, \{e_{25}, e_{26}, e_{29}, e_{30}\}$).

Due to that it is only Tests 3 and 4 that are different in MCDS and ICDS, only these two tests are considered in the simulation study.

6.2 MODEL USED IN THE DIAGNOSIS SYSTEM

The model of the vehicle powertrain used in the diagnosis systems is the same as the model presented in Section 3.1 and in Appendix A, except for the clutch model. To investigate the consequences of not having a valid model in all operating modes of the system to be monitored, it is assumed that the model of the clutch only is available when the clutch is fully engaged. This results in that when the clutch is disengaged or there is slip in the clutch, corresponding test quantities are not updated and no faults are to be detected in these tests.

6.3 INITIALIZATION OF STATES WHEN RESTARTING RESIDUAL GENERATORS

The time it takes for a transient in a dynamic residual generator to fade out after it is initialized decreases if the states in the residual are accurately initialized. When the model is not valid in all operating points it is therefore more important to accurately initialize the states since the residual is restarted when the model becomes valid. The basic idea when initializing the states is to use (24). There are several possibilities to reduce the sensor noise impact on the initialization of the state. In e.g. Krysanter et al. (2010) this is done by assuming Gaussian noise and finding the initial value of the state using a least square estimate over a time window. A less complex method is to filter the signal to be used in the initialization using a time constant, τ . This method is used here, and the state w_w calculated from (35) and (11) in ICDS is reinitialized in the residual

generators when the vehicle model used in the diagnostic tests is becoming valid using $r(t_0) = 0$ in (36)

$$\omega_w(t_0) = \frac{1}{\tau_w p + 1} \frac{\omega_{gb,sens}(t)}{u_f} \Big|_{t=t_0} \quad (37)$$

The fuel consumed, m_f , is a state in addition to ω_w in Test 3, while there are two additional states in Test 4; m_f and SoC . The states except ω_w have slow dynamics and are therefore not reinitialized when the diagnostic model becomes valid. Instead the states are only updated when the model is valid.

The state used in the transformation in MCDS is reinitialized when the model used in the diagnosis system is getting valid by using (25), where $x_i = \omega_{gb}$ according to (32). It is assumed that the powertrain is fault free in the initialization of Γ , i.e. $r(t_0) = 0$, and $\omega_{gb,sens}$ is used instead of ω_{gb}

$$\Gamma(t_0) = - \frac{1}{\tau_\Gamma p + 1} \alpha \omega_{gb,sens}(t) \Big|_{t=t_0} \quad (38)$$

6.4 TWO APPROACHES FOR WHEN TO UPDATE DYNAMIC TEST QUANTITIES

As stated above, it is assumed that the monitored system is fault free and the residual is zero in the initialization of the states in both MCDS and ICDS. If the equations used in the expression for the signal to be integrated are inconsistent with the monitored system, the integrated signal will drift from the true value. To increase the fault sensitivity of corresponding test quantities, it is preferable to only update the test quantities when the residuals are non-zero even though the estimation of the signal to be integrated is inconsistent (see Figure 4). Two approaches to find updating conditions for the test quantities are presented below.

FIXED TIME

The first approach is to not update the dynamic test quantities in the diagnosis system before a time, t_d , after a test has been valid in order for the fault to have time to affect r . This means that the test quantity is updated when

$$t > t_0 + t_d \quad (39)$$

where t_0 is the time of the latest reinitialization of the states.

INTERNAL FORM

The second approach requires fault models that are used to investigate how the faults affect the residuals by finding the internal form of the residual generators, see Section 5.3. This approach is implemented and compared with the fixed

time approach given above using Test 4 in ICDS. The test is expected to react on $f_{em,\eta}$, $f_{b,sc}$, and $f_{em,U,sens}$, but for simplicity only $f_{em,\eta}$ and $f_{em,U,sens}$ are considered here. These faults lead to different internal forms of the residual generator and therefore several test quantities are constructed that use different updating conditions.

First, the internal form of Test 4 in ICDS when there is a fault in the voltage sensor in the electric machine is found by the substitution chain given by the computation sequence (50)

$$r(t) = \frac{f_{em,U,sens}}{f_{em,U,sens} + 1} \cdot C \int_{t_0}^t u_{gb} I_{em}(\tau) \eta_{em,0}^{\text{sign}\{I_{em}(\tau)\}} d\tau \quad (40)$$

where C is a constant and u_{gb} the gear ratio in the gearbox. Due to (40) it is obvious that the fault excitation is dependent on the magnitude of the integral. This leads to that a condition for when to update the test quantity to achieve good fault detection performance is

$$\left| \int_{t_0}^t u_{gb}(\tau) I_{em}(\tau) \eta_{em,0}^{\text{sign}\{I_{em}(\tau)\}} d\tau \right| > J_1 \quad (41)$$

where J_1 is a design parameter. A comparison of the conditions in (39) and (41) is shown in Figure 5. As expected, the test quantity that is updated using (41) does not decrease in the second time interval the model is valid and there is low fault excitation in the residual. Note that there is low fault excitation in the residual at this time interval even though the electric machine is used. This is due to that the machine frequently switches from generator to motor mode, see Figure 5.

The internal form of the residual generator for $f_{em,\eta}$

$$r(t) = C \int_{t_0}^t u_{gb}(\tau) I_{em}(\tau) \eta_{em,0}^{\text{sign}\{I_{em}(\tau)\}} \cdot \left((1 + f_{em,\eta})^{\text{sign}\{I_{em}(\tau)\}} - 1 \right) d\tau \quad (42)$$

Finding the times $\tau \in \{t_0, t\}$ when $I_{em}(t) \geq 0A$ and $I_{em}(t) < 0A$

$$t^+ = \{\tau \in \{t_0, t\} : I_{em}(t) \geq 0A\} \quad (43a)$$

$$t^- = \{\tau \in \{t_0, t\} : I_{em}(t) < 0A\} \quad (43b)$$

leads to that (42) can be written as

$$r(t) = f_{em,\eta} \cdot C \int_{t^+} u_{gb} I_{em}(\tau) \eta_{em,0} d\tau + \frac{f_{em,\eta}}{1 + f_{em,\eta}} \cdot C \int_{t^-} (-) u_{gb} \frac{I_{em}(\tau)}{\eta_{em,0}} d\tau \quad (44)$$

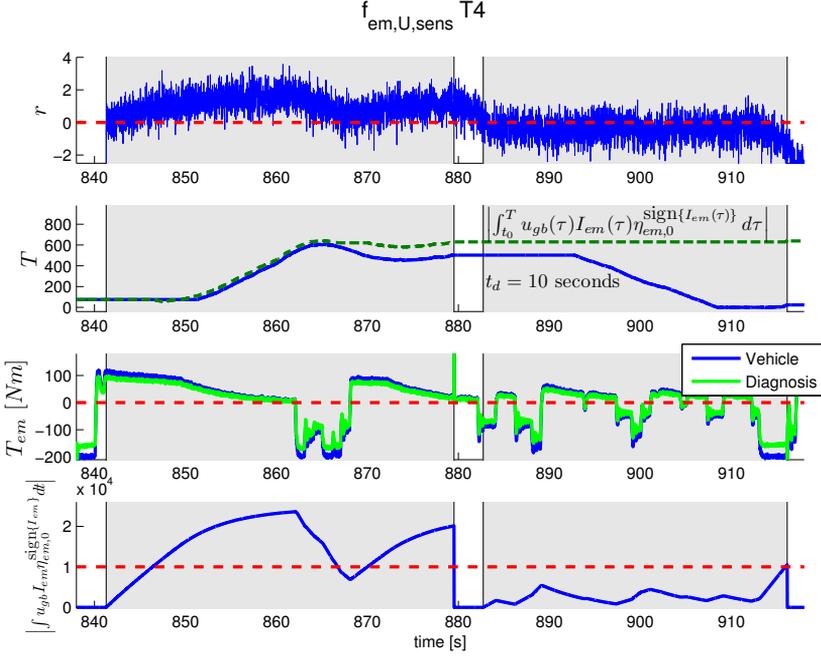


Figure 5: The residual and test quantities for Test 4 in ICDS when $f_{em,U,sens}$ has occurred, and the shaded areas indicate when the model used in the diagnosis system is valid. Two different conditions, (39) and (41), for when the test quantity is to be updated are compared, using $t_d = 10s$ and $J_1 = 10000As$. The torque T_{em} is almost proportional to I_{em} , see (8b), and the operating points of the electric machine thereby affect how long time the test quantity is updated when (41) is used as the update condition.

Under the condition that $f_{em,\eta} \in]-1, 0]$, which is a reasonable assumption according to (8c) and (13c), $\dot{r} < 0$ for both positive and negative I_{em} . This leads to the following condition for when to update the test quantity

$$\int_{t_0}^t u_{gb} \left| I_{em}(\tau) \eta_{em,0}^{\text{sign}\{I_{em}(\tau)\}} \right| d\tau > J_2 \quad (45)$$

where J_2 is a design parameter. The difference between (41) and (45) is that the latter test quantity is updated as long as the electric machine is used, even though it changes operation mode between generator and motor. Figure 6 presents the test quantities achieved using the updating conditions in (39), (41), and (45) when $f_{em,\eta}$ has occurred. It can be seen that the test quantity updated when (45) is true reacts better than the test quantity updated when (41) is used, since the first is updated in both time intervals the model is valid and the

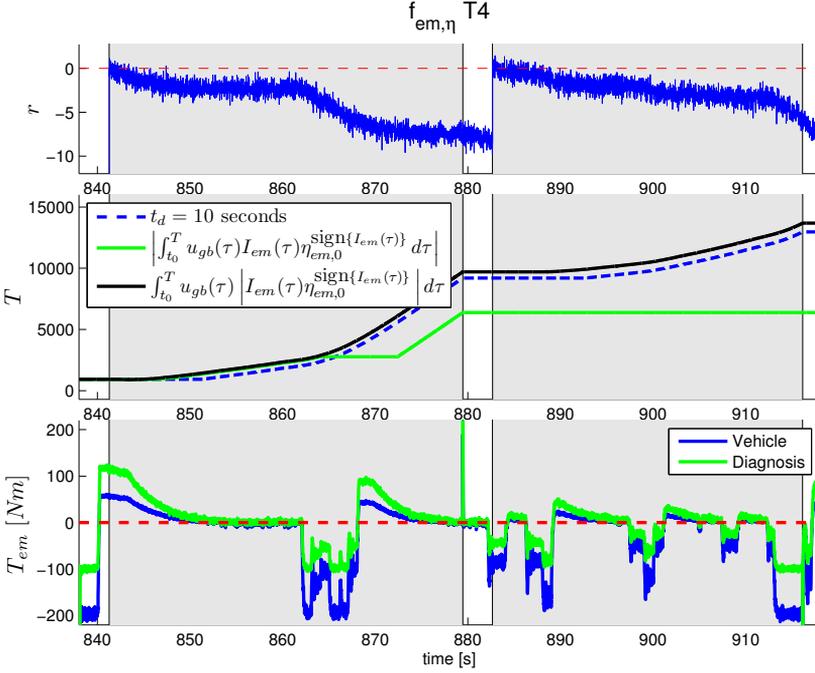


Figure 6: Test quantities for Test 4 in ICDS based on the same residual but different updating conditions, when $f_{em,\eta}$ has occurred. The test quantities are updated with the conditions given by (39), (41) and (45). In contradiction to when $f_{em,U,sens}$ has occurred (see Figure 5), $\dot{r} < 0$ both when the machine operates as motor and generator, as indicated in (44).

specific fault excite the residual at these times.

Remark 4 (Several test quantities based on the same residual). *Based on Figures 5 and 6, two different conditions are derived for when to update the test quantity, i.e. (41) and (45), in order for Test 4 to react on both $f_{em,U,sens}$ and $f_{em,\eta}$ in a good way. Thus, it makes sense to compute several test quantities based on the same residual. The drawback is the increase in computational complexity.*

6.5 SIMULATIONS ON DRIVING CYCLE

A simulation study is carried out to evaluate the designed diagnosis systems. The faults described in (13) are induced in the vehicle model one by one. The driving cycle used is FTP75 and the speed profile can be found in e.g. Guzzella and Sciarretta (2007).

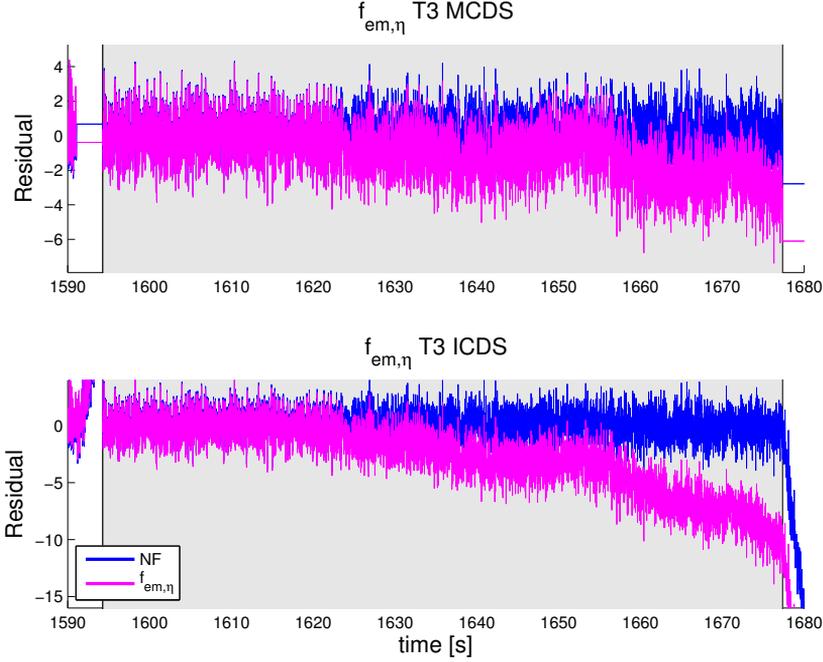


Figure 7: The residual used in Test 3 in ICDS is more affected to $f_{em,\eta}$ compared to corresponding residual in MCDS. The residuals are normalized so the standard deviation in the fault free case is one, and the shaded areas indicate when the model used in the diagnosis system is valid, i.e. from 1594 seconds to 1677 seconds.

NORMALIZED TEST QUANTITIES

Test quantities are computed using the CUSUM algorithm presented in Section 5.3. To compare the performance of different diagnosis systems, normalized test quantities, T_{norm} , are calculated based on the maximum value, $T_{\text{max,NF}}$, of T in the fault free case when FTP75 is used. An alarm is generated when $T_{\text{norm}} > 1$ and the design parameter $\Phi \in [0, 1]$ states the margin to false alarm for the specific driving cycle.

$$T_{\text{norm}} = \frac{T}{T_{\text{max,NF}}} \Phi \quad (46)$$

RESIDUAL RESPONSES

Residuals for the diagnosis systems are compared, and in Figure 7 the residuals used in Tests 3 in MCDS and ICDS are shown for the fault free case as well as when $f_{em,\eta}$ has occurred. The residuals are normalized so the standard deviation

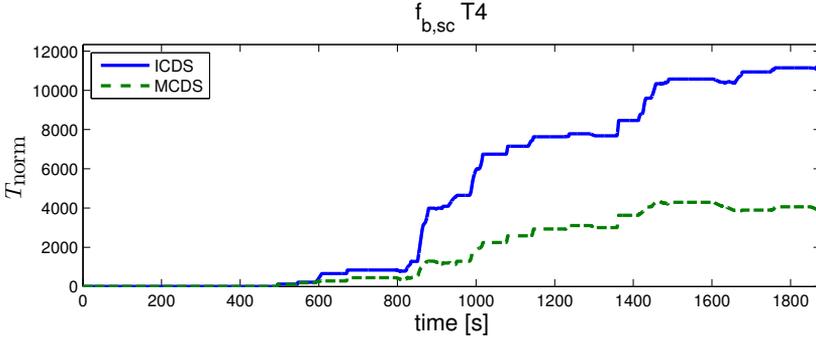


Figure 8: Normalized test quantities for Test 4 in MCDS and ICDS when there is a short circuit in the battery after 400 seconds. The test quantities are normalized according to (46) with $\Phi = 0.5$, and T_{norm} using integral causality is larger than T_{norm} using mixed causality after the fault has occurred.

of the residuals in the fault free case is one. The residual used in ICDS is more affected by the fault compared to the residual in MCDS, and this is the general result in the simulation study when the different faults are induced in the vehicle model.

The method used to reinitialize the states in the residuals when the model is getting valid is working properly, since the values of the residuals are close to zero after the initialization.

TEST QUANTITY RESPONSES

The residuals are post processed as described above, and all faults are isolated in both MCDS and ICDS. The test quantities generally react stronger in ICDS even though the tests in the two systems are based on the same sets of model equations. In Figure 8 the test quantities are shown when there is a short circuit in the battery after 400 seconds.

The different updating conditions for the test quantities are implemented in ICDS to evaluate the impact of these different design selections using the entire driving cycle. As indicated in Figures 5 and 6, the tests generally react stronger on the fault if the internal form of the residual is considered when designing test quantities. In Figure 9 this is exemplified when $f_{em,U,sens}$ has occurred and the test quantity that is updated using the condition based on the internal form for the specific fault, i.e. (41), reacts better than the other test quantities.

6.6 SUMMING UP

The method used to initialize the states in the diagnosis systems is straight forward and is working properly, i.e. the residuals are close to zero after the

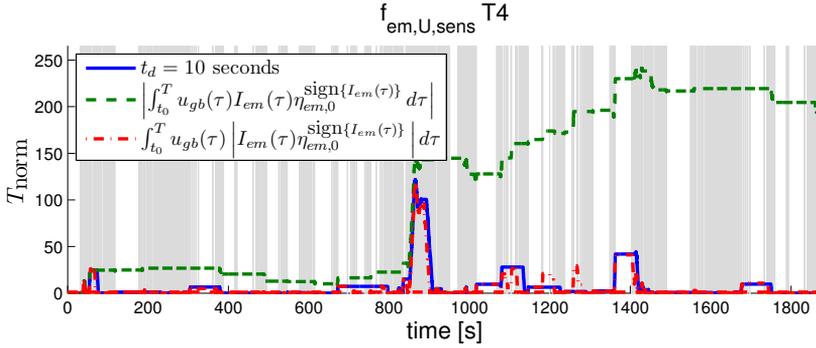


Figure 9: The test quantities for Test 4 in ICDS using different conditions for when to update the test quantities and there is a fault in the voltage sensor in the electric machine at $t = 0s$. The shaded areas indicate when the model used in the diagnosis system is valid. At $t = 840s$ the system is excited and the diagnostic tests react on the fault (see Figure 5). Note that it is only the test quantity that is updated using the internal form of the fault, i.e. (41), that reacts from this time till the end of the driving cycle.

initialization. The use of the internal form is demonstrated to be advantageous, according to (41), (45) and Figures 5, 6, and 9. Note that in order for Test 4 to react on both $f_{em,U,sens}$ and $f_{em,\eta}$ in a good way, two test quantities based on the same residual using different updating conditions are used, since the faults affect the residual in different ways. Since the uncertainties in the study predominantly relate to disturbances rather than model inaccuracies, then it is reasonable that ICDS performs better than MCDS. Therefore it is valuable to use the modified algorithm FINDRESIDUALGENERATORSMOD that singles out such systems.

7 CONCLUSIONS

A reduced HEV model was used to illustrate how the selection of consistency relation affects the occurrence of unique expressions of the unknowns and algebraic loops in sequential residual generators. All sequential residual generator candidates of the vehicle model were investigated with respect to algebraic and dynamic properties, and it was found that in general there are few selections of consistency relations in an MSO that achieves predefined properties. Thereby it can be stated that systematic methods analyzing the properties of the residual generator candidates are valuable.

Having selected the residuals, important aspects remain when utilizing them in a diagnosis system. A straight forward method, given in (24) and e.g. (37),

to reinitialize the states in the residual generators was shown to work properly. It was also shown that the method to update the test quantities based on the internal form of the residual generators, significantly increased the diagnosis performance. Simulations verified that it is beneficial to use several test quantities based on the same residual using different updating conditions when a residual is sensitive for several faults.

All in all, it has been shown that the engineering support the used methods gives was a key to design well behaved diagnosis systems. The methods are general in character and provides a useful methodology when designing diagnosis systems for HEVs or other complex systems.

REFERENCES

- L. Barford, E. Manders, G. Biswas, P. Mosterman, V. Ram, and J. Barnett. Derivative estimation for diagnosis. Technical report, HP Laboratories Palo Alto, 1999. HPL-1999-18.
- M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer, 2nd edition, 2006.
- K. Chau, C. Chan, and C. Liu. Overview of permanent-magnet brushless drives for electric and hybrid electric vehicles. *Industrial Electronics, IEEE Trans. on*, 55(6):2246–2257, June 2008.
- E. Chow and A. Willsky. Analytical redundancy and the design of robust failure detection systems. *Automatic Control, IEEE Transactions on*, 29(7):603–614, July 1984.
- M.-O. Cordier, P. Dague, F. Levy, J. Montmain, M. Staroswiecki, and L. Trave-Massuyes. Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspectives. *IEEE Trans. on SMC – Part B*, 34(5):2163–2177, October 2004.
- L. Eriksson. Simulation of a vehicle in longitudinal motion with clutch engagement and release. In *IFAC Workshop: Advances in Automotive Control*, pages 65–70, Karlsruhe, Germany, 2001.
- P. M. Frank. Enhancement of robustness in observer-based fault detection. *International J. of Control*, 59(4):955–981, 1994.
- J. Fredriksson, J. Larsson, J. Sjöberg, and P. Krus. Evaluating hybrid electric and fuel cell vehicles using the capsim simulation environment. In *22nd International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exposition*, pages 1994–2004, Yokohama, Japan, 2006.
- E. Frisk and J. Åslund. Lowering orders of derivatives in non-linear residual generation using realization theory. *Automatica*, 41(10):1799–1807, 2005.
- E. Frisk and M. Nyberg. A minimal polynomial basis solution to residual generation for fault diagnosis in linear systems. *Automatica*, 37(9):1417–1424, September 2001.
- E. Frisk, A. Bregon, J. Åslund, M. Krysander, B. Pulido, and G. Biswas. Diagnosability analysis considering causal interpretations for differential constraints. *IEEE Trans. on SMC – Part A: Systems and Humans*, 42(5):1216–1229, September 2012.
- S. Gentil, J. Montmain, and C. Combastel. Combining FDI and AI approaches within causal-model-based diagnosis. *IEEE Trans. on SMC – Part B*, 34(5):2207–2221, October 2004.

- F. Gustafsson. *Adaptive filtering and change detection*. John Wiley & Sons, 2000.
- L. Guzzella and A. Amstutz. CAE tools for quasi-static modeling and optimization of hybrid powertrains. *IEEE Trans. on Vehicular Technology*, 48(6): 1762–1769, November 1999.
- L. Guzzella and A. Sciarretta. *Vehicle Propulsion System, Introduction to Modeling and Optimization*. Springer Verlag, Zürich, 2 edition, 2007.
- A. R. Hambley. *Electrical Engineering, principles and applications*. Pearson Education, Upper Sadle River, 3 edition, 2005.
- R. Isermann. *Fault Diagnosis Systems - An Introduction from fault Detection to Fault Tolerance*. Springer Verlag, 2006.
- G. Katsillis and M. Chantler. Can dependency-based diagnosis cope with simultaneous equations? In *8th International Workshop on Principles of Diagnosis (DX-97)*, pages 51–59, Le Moint Saint Michel, France, 1997.
- M. Krysander and E. Frisk. Sensor placement for fault diagnosis. *IEEE Trans. on SMC – Part A*, 38(6):1398–1410, 2008.
- M. Krysander, J. Åslund, and M. Nyberg. An efficient algorithm for finding minimal over-constrained sub-systems for model-based diagnosis. *IEEE Trans. on SMC – Part A*, 38(1), 2008.
- M. Krysander, F. Heintz, J. Roll, and E. Frisk. FlexDx: A reconfigurable diagnosis framework. *Engineering Applications of Artificial Intelligence*, 23(8): 1303–1313, October 2010.
- P. Nelson, K. Amine, A. Rousseau, and H. Yomoto. Advanced lithium-ion batteries for plug-in hybrid-electric vehicles. In *Proceedings in 23rd International Electric Vehicle Symposium*, pages 1589–1605, Anaheim, CA, 2007.
- E. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.
- T. B. Reddy. *Linden's Handbook of Batteries*. McGraw-Hill, 2011.
- G. Rizzoni, L. Guzzella, and B. Baumann. Unified modeling of hybrid electric vehicle drivetrains. *Mechatronics, IEEE/ASME Transactions on*, 4(3):246–257, September 1999.
- M. Staroswiecki and G. Comtet-Varga. Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems. *Automatica*, 37(5): 687–699, 2001.
- M. Staroswiecki and P. Declerck. Analytical redundancy in non-linear interconnected systems by means of structural analysis. In *Proceedings of IFAC AIPAC'89*, pages 51–55, Nancy, France, 1989.

- C. Sundström. Vehicle level diagnosis for hybrid powertrains. Technical report, 2011. Licentiate thesis LiU-TEK-LIC-2011:27, Thesis No. 1488.
- C. Sundström, E. Frisk, and L. Nielsen. Overall monitoring and diagnosis for hybrid vehicle powertrains. In *6th IFAC Symposium on Advances in Automotive Control*, pages 119–124, Munich, Germany, 2010.
- C. Svärd and M. Nyberg. Residual generators for fault diagnosis using computation sequences with mixed causality applied to automotive systems. *IEEE Trans. on SMC – Part A*, 40(6):1310–1328, 2010.

A POWERTRAIN MODEL

$$\begin{aligned}
e1: \quad & \dot{m}_f = \text{ice}_{ctrl} \frac{\omega_e}{4\pi q_{LHV}} \\
e2: \quad & T_e = \left(\text{ice}_{ctrl} \frac{4\eta_{e,i}}{N_{cyl} \pi S B^2} - p_{me0,f} - p_{me0,g} \right) N_{cyl} \frac{S B^2}{16} \\
e3: \quad & p_{me0,f} = k_1 (k_2 + k_3 S^2 \omega_e^2) \Pi_{bl} \sqrt{\frac{k_4}{B}} \\
e4: \quad & \dot{SoC} = -\frac{1}{Q_b} I_b \\
e5: \quad & U_{oc} = f_1(SoC) \\
e6: \quad & U_b = U_{oc} - R_b I_b \\
e7: \quad & U_{em} = U_{em,ctrl} \\
e8: \quad & I_{em} = \frac{U_{em} - k_i \omega_{em}}{R_{em}} \\
e9: \quad & k_a = k_i \eta_{em,0}^{\text{sign}\{I_{em}\}} \\
e10: \quad & T_{em} = I_{em} k_a - c_{em,f} \omega_{em} \\
e11: \quad & I_b = \frac{I_{em} U_{em}}{U_b} \\
e12: \quad & T_c = u_c T_e \\
e13: \quad & \omega_e = \begin{cases} \omega_{e,idle}, & \omega_{mj} < \omega_{e,idle} \\ \omega_{mj}, & \omega_{mj} \geq \omega_{e,idle} \end{cases} \\
e14: \quad & T_{mj} = T_{em} u_{em} + T_c \\
e15: \quad & J_{mj} = J_{em} u_{em}^2 + J_c + J_e \\
e16: \quad & \omega_{em} = \frac{1}{u_{em}} \omega_{mj} \\
e17: \quad & J_{gb} = f_2(\text{gear}) \\
e18: \quad & u_{gb} = f_3(\text{gear}) \\
e19: \quad & \eta_{gb} = \begin{cases} \eta_{pos}, & T_{mj} > T_{gb,l} \\ \eta_{neg}, & T_{mj} \leq T_{gb,l} \end{cases} \\
e20: \quad & T_{gb,l} = f_4(\text{gear}, \omega_e) \\
e21: \quad & T_{gb} = (T_{mj} - T_{gb,l}) \eta_{gb} u_{gb} \\
e22: \quad & J_{tot} = u_{gb}^2 (J_{gb} + J_{mj}) \\
e23: \quad & \omega_{mj} = u_{gb} \omega_{gb} \\
e24: \quad & m_v = m_{v,0} - m_f \\
e25: \quad & T_d = \frac{1}{2} \rho C_d A_f \omega_w^2 r_w^3 \\
e26: \quad & T_r = \begin{cases} m_v g C_r r_w, & 1000 \omega_w > m_v g C_r r_w \\ 1000 \omega_w, & -m_v g C_r r_w \leq 1000 \omega_w < m_v g C_r r_w \\ -m_v g C_r r_w, & 1000 \omega_w \leq -m_v g C_r r_w \end{cases} \\
e27: \quad & T_g = m_v g r_w \sin \theta \\
e28: \quad & T_b = T_{b,ctrl} \\
e29: \quad & T_{net} = T_{gb} u_f - T_d - T_b - T_r - T_g
\end{aligned}$$

$$\begin{aligned}
e30: \quad \dot{\omega}_w &= \frac{T_{net}}{J_{tot}u_f^2 + m_v r_w^2} \\
e31: \quad v &= \omega_w r_w \\
e32: \quad \omega_{gb} &= \omega_w u_f \\
e33: \quad I_{b,sens} &= I_b \\
e34: \quad U_{em,sens} &= U_{em} \\
e35: \quad I_{em,sens} &= I_{em} \\
e36: \quad \omega_{e,sens} &= \omega_e \\
e37: \quad \omega_{gb,sens} &= \omega_{gb}
\end{aligned}$$

B RESIDUAL GENERATORS

B.1 SAME TESTS IN ICDS AND MCDS

Test 1:

$$\mathcal{C}_{T1} = \{(\{U_{em}\}, \{e_7\})\} \quad (47a)$$

$$ARR_{T1} = e_{34} \quad (47b)$$

Test 2:

$$\begin{aligned}
\mathcal{C}_{T2} = & \{(\{I_b\}, \{e_{33}\}), (\{I_{em}\}, \{e_{35}\}), (\{\omega_{gb}\}, \{e_{37}\}), \\
& (\{u_{gb}\}, \{e_{18}\}), (\{\omega_{mj}\}, \{e_{23}\}), (\{\omega_{em}\}, \{e_{16}\}), (\{SoC\}, \{e_4\}), \\
& (\{U_{oc}\}, \{e_5\}), (\{U_b\}, \{e_6\}), (\{U_{em}\}, \{e_8\})\} \quad (48a)
\end{aligned}$$

$$ARR_{T2} = e_{11} \quad (48b)$$

B.2 ICDS

Test 3:

$$\begin{aligned}
\mathcal{C}_{T3,I} = & \{(\{\omega_e\}, \{e_{36}\}), (\{\omega_{gb}\}, \{e_{37}\}), (\{u_{gb}\}, \{e_{18}\}), \\
& (\{\omega_{mj}\}, \{e_{23}\}), (\{\omega_{em}\}, \{e_{16}\}), (\{U_{em}\}, \{e_7\}), (\{I_{em}\}, \{e_8\}), \\
& (\{k_a\}, \{e_9\}), (\{T_{em}\}, \{e_{10}\}), (\{p_{me0,f}\}, \{e_3\}), (\{T_e\}, \{e_2\}), \\
& (\{m_f\}, \{e_1\}), (\{J_{gb}\}, \{e_{17}\}), (\{J_{mj}\}, \{e_{15}\}), (\{J_{tot}\}, \{e_{22}\}), \\
& (\{T_{gb,l}\}, \{e_{20}\}), (\{T_c\}, \{e_{12}\}), (\{T_{mj}\}, \{e_{14}\}), (\{\eta_{gb}\}, \{e_{19}\}), \\
& (\{m_v\}, \{e_{24}\}), (\{T_g\}, \{e_{27}\}), (\{T_{gb}\}, \{e_{21}\}), (\{T_b\}, \{e_{28}\}), \\
& (\{T_d, T_r, T_{net}, \omega_w\}, \{e_{25}, e_{26}, e_{29}, e_{30}\})\} \quad (49a)
\end{aligned}$$

$$ARR_{T3,I} = e_{32} \quad (49b)$$

Test 4:

$$\begin{aligned}
\mathcal{C}_{T4,I} = & (\{I_b\}, \{e_{33}\}), \{(\{U_{em}\}, \{e_{34}\}), (\{\omega_e\}, \{e_{36}\}), \\
& (\{\omega_{gb}\}, \{e_{37}\}), (\{u_{gb}\}, \{e_{18}\}), (\{\omega_{mj}\}, \{e_{23}\}), (\{\omega_{em}\}, \{e_{16}\}), \\
& (\{SoC\}, \{e_4\}), (\{U_{oc}\}, \{e_5\}), (\{U_b\}, \{e_6\}), (\{I_{em}\}, \{e_{11}\}), \\
& (\{k_a\}, \{e_9\}), (\{T_{em}\}, \{e_{10}\}), (\{p_{me0,f}\}, \{e_3\}), (\{T_e\}, \{e_2\}), \\
& (\{m_f\}, \{e_1\}), (\{J_{gb}\}, \{e_{17}\}), (\{J_{mj}\}, \{e_{15}\}), (\{J_{tot}\}, \{e_{22}\}), \\
& (\{T_{gb,l}\}, \{e_{20}\}), (\{T_c\}, \{e_{12}\}), (\{T_{mj}\}, \{e_{14}\}), (\{\eta_{gb}\}, \{e_{19}\}), \\
& (\{m_v\}, \{e_{24}\}), (\{T_g\}, \{e_{27}\}), (\{T_{gb}\}, \{e_{21}\}), (\{T_b\}, \{e_{28}\}), \\
& (\{T_d, T_r, T_{net}, \omega_w\}, \{e_{25}, e_{26}, e_{29}, e_{30}\}) \} \quad (50a)
\end{aligned}$$

$$ARR_{T4,I} = e_{32} \quad (50b)$$

B.3 MCDS

Test 3:

$$\begin{aligned}
\mathcal{C}_{T3,M} = & \{(\{\omega_e\}, \{e_{36}\}), (\{\omega_{gb}\}, \{e_{37}\}), (\{u_{gb}\}, \{e_{18}\}), \\
& (\{\omega_{mj}\}, \{e_{23}\}), (\{\omega_{em}\}, \{e_{16}\}), (\{U_{em}\}, \{e_7\}), (\{I_{em}\}, \{e_8\}), \\
& (\{k_a\}, \{e_9\}), (\{T_{em}\}, \{e_{10}\}), (\{p_{me0,f}\}, \{e_3\}), (\{T_e\}, \{e_2\}), \\
& (\{m_f\}, \{e_1\}), (\{J_{gb}\}, \{e_{17}\}), (\{J_{mj}\}, \{e_{15}\}), (\{J_{tot}\}, \{e_{22}\}), \\
& (\{T_{gb,l}\}, \{e_{20}\}), (\{T_c\}, \{e_{12}\}), (\{T_{mj}\}, \{e_{14}\}), (\{\eta_{gb}\}, \{e_{19}\}), \\
& (\{m_v\}, \{e_{24}\}), (\{T_g\}, \{e_{27}\}), (\{T_{gb}\}, \{e_{21}\}), (\{T_b\}, \{e_{28}\}), \\
& (\{\omega_w\}, \{e_{32}\}), (\{T_d\}, \{e_{25}\}), (\{T_r\}, \{e_{26}\}), (\{T_{net}\}, \{e_{29}\}) \} \quad (51a)
\end{aligned}$$

$$ARR_{T3,M} = e_{30} \quad (51b)$$

Test 4:

$$\begin{aligned}
\mathcal{C}_{T4,M} = & (\{I_b\}, \{e_{33}\}), \{(\{U_{em}\}, \{e_{34}\}), (\{\omega_e\}, \{e_{36}\}), \\
& (\{\omega_{gb}\}, \{e_{37}\}), (\{u_{gb}\}, \{e_{18}\}), (\{\omega_{mj}\}, \{e_{23}\}), (\{\omega_{em}\}, \{e_{16}\}), \\
& (\{SoC\}, \{e_4\}), (\{U_{oc}\}, \{e_5\}), (\{U_b\}, \{e_6\}), (\{I_{em}\}, \{e_{11}\}), \\
& (\{k_a\}, \{e_9\}), (\{T_{em}\}, \{e_{10}\}), (\{p_{me0,f}\}, \{e_3\}), (\{T_e\}, \{e_2\}), \\
& (\{m_f\}, \{e_1\}), (\{J_{gb}\}, \{e_{17}\}), (\{J_{mj}\}, \{e_{15}\}), (\{J_{tot}\}, \{e_{22}\}), \\
& (\{T_{gb,l}\}, \{e_{20}\}), (\{T_c\}, \{e_{12}\}), (\{T_{mj}\}, \{e_{14}\}), (\{\eta_{gb}\}, \{e_{19}\}), \\
& (\{m_v\}, \{e_{24}\}), (\{T_g\}, \{e_{27}\}), (\{T_{gb}\}, \{e_{21}\}), (\{T_b\}, \{e_{28}\}), \\
& (\{\omega_w\}, \{e_{32}\}), (\{T_d\}, \{e_{25}\}), (\{T_r\}, \{e_{26}\}), (\{T_{net}\}, \{e_{29}\}) \} \quad (52a)
\end{aligned}$$

$$ARR_{T4,M} = e_{30} \quad (52b)$$

Diagnostic Method Combining Map and Fault
Models Applied on a Hybrid Electric Vehicle*

D

*Submitted to Journal.

Diagnostic Method Combining Map and Fault Models Applied on a Hybrid Electric Vehicle

Christofer Sundström, Erik Frisk, and Lars Nielsen

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, SE-581 83 Linköping, Sweden.*

ABSTRACT

A common situation in the automotive industry is that map based models are available. In general these models accurately describe the fault free system, and are therefore suited for fault detectability in a diagnosis system. However, one drawback using such a model is that fault isolation then requires that measurements of the faulty system is done, which is costly. Another approach is to use a model of the system where the faults are explicitly included. To directly achieve good diagnostic performance such a model needs to be accurate, which also is costly. Therefore, in the new approach taken here, two models are used in combination to achieve both good fault detectability and isolability in a diagnosis system; one is a map based model, and one is describing how the faults affect the system. The approach is exemplified by designing a diagnosis system monitoring the power electronics and the electric machine in a hybrid electric vehicle. In an extensive simulation study it is shown that the approach works well and is a promising path to achieve both good fault detectability and isolability performance, without the need for neither measurements of a faulty system nor detailed physical modeling. In the designed diagnosis system all faults are fully isolated, and the size of the faults are accurately estimated.

1 INTRODUCTION

Fault monitoring and diagnosis is used to detect and isolate faults in a system. Several approaches can be used, and one common is consistency based diagnosis de Kleer et al. (1992) using residual generators Blanke et al. (2006). Such diagnosis systems compare the consistency between observations and a model of the system to be monitored. The models are developed to different level of detail, and a common approach in the automotive industry is map based models directly calibrated from measurements. These models are straightforward to design using measurements, and one benefit of using such a map based model is that it accurately describes the outputs. The high model accuracy directly results in good fault detection performance, but one drawback with a map based model is the difficulty to isolate faults from each other, since internal physical phenomena are not described by the model. One way to achieve fault isolability using a map based model is to collect data when the faults have occurred in the system to be monitored, which is a costly solution due to the many fault cases. Another approach to achieve good fault isolability is to use models that explicitly describes how different faults affect the system to be monitored. To achieve good diagnostic performance using only such a model, it needs to be accurate including detailed physical modeling, which also is costly.

1.1 CONTRIBUTIONS AND OUTLINE

The main idea here is an approach to combine two models, which means using a map describing the fault free system in combination with a model describing how the faults affect the system. A preliminary version of this idea is presented in Sundström et al. (2013). Compared to that paper, the concepts have been clarified and generalized, but most important is that analysis and an extensive simulation study show the importance and performance of the diagnostic method. It is demonstrated that the benefit of the proposed approach is that measurements of a faulty system is not needed, and that the accuracy demands on the model used for fault modeling are lower than for designing a diagnosis system without using the map based model.

The proposed approach is used in the design of a diagnosis system monitoring the power electronics and the electric machine used in a hybrid electric vehicle (HEV), where monitoring of the components is important in order to achieve high up-time of the vehicle. Further issues are safety and component protection, especially the battery that is sensitive and costly Chen et al. (2013). The models used in the diagnosis design of the electric machine are described in Section 2, and in Section 3 these models are combined to include fault models in the map based model. The value of using this combined model in the diagnosis system is evaluated in Section 5, and finally the conclusions are given in Section 6.

2 MODELS OF THE ELECTRIC MACHINE

In HEVs mainly permanent magnet synchronous machines (PMSM) are used since this type of machine in general has higher efficiency compared to other machine types Zhu and Howe (2007); Mellor (1999). A PMSM is an AC machine, but it is possible to use a DC source, e.g. a battery, and use power electronics to achieve an alternating current.

Two models of a PMSM are presented, that later are used to illustrate the approach in the design of the diagnosis system. The first model includes a map that describes the power losses in the machine and is presented in Section 2.1. To model how faults affect the machine and power electronics, which is not captured in the map based model, the second model is based on analytical expressions and is presented in Section 2.2.

2.1 MAP BASED MODEL

The map based model describe the power losses in the machine and the power electronics, and is based on measurements to find the difference between the electrical and mechanical powers. The map of the power losses, $P_{em,l}^{map}$, is three dimensional taking the delivered torque, T_{em} , motor speed, ω_{em} , and battery voltage, U_b , as inputs

$$P_{em,l}^{map} = f(T_{em}, \omega_{em}, U_b) \quad (1)$$

and the power losses are given in Figure 1. There are limitations in the delivered torque from the machine, denoted $T_{em,min}$ in generator mode and $T_{em,max}$ in motor mode, that are functions of ω_{em} and U_b . The limited torque, $T_{em,lim}$, is equal to the requested torque, $T_{em,req}$, if the requested torque is within the limitations of what the machine is able to deliver

$$T_{em,lim} = \begin{cases} T_{em,min}, & T_{em,req} < T_{em,min} \\ T_{em,req}, & T_{em,min} \leq T_{em,req} < T_{em,max} \\ T_{em,max}, & T_{em,req} \geq T_{em,max} \end{cases} \quad (2)$$

The delivered torque is computed by filtering $T_{em,lim}$

$$T_{em} = \frac{1}{\tau_{em}s + 1} T_{em,lim} \quad (3)$$

and the mechanical power delivered by the machine

$$P_{em,m} = T_{em}\omega_{em} \quad (4)$$

is used to calculate the electrical power

$$P_{em,e} = P_{em,m} + P_{em,l}^{map} \quad (5)$$

The power electronics is included in the model and is assumed to be an ideal component. The battery current, I_b , is computed by dividing $P_{em,e}$ with the battery terminal voltage, U_b

$$I_b = \frac{P_{em,e}}{U_b} \quad (6)$$

2.2 ANALYTICAL MODEL

A PMSM can be modeled as a separately excited DC motor with constant field Guzzella and Sciarretta (2007), since the stator of a PMSM consists of windings, and the armature of permanent magnets. This is done in the model based on analytical expressions, where the resistive and frictional losses are modeled to represent the losses of the machine. The torque T_{em} is modeled to be proportional to the current, I_{em} , except for the frictional losses that are modeled to be proportional to ω_{em} Zhu et al. (2000). The output torque from the machine is

$$T_{em} = kI_{em} - c_f\omega_{em} \quad (7)$$

where c_f is a friction constant and k is a machine constant. The current is calculated using the voltage, U_{em} , supplied by the power electronics and the electromotive force (emf), that depends on the speed of the machine, ω_{em}

$$I_{em} = \frac{1}{R_{em}} (U_{em} - \underbrace{k\omega_{em}}_{emf}) \quad (8)$$

where R_{em} is the resistance in the electric machine. The power losses in the machine are computed using

$$P_{em,l}^a = U_{em}I_{em} - T_{em}\omega_{em} \quad (9)$$

Substituting U_{em} and I_{em} from (7) and (8) gives

$$P_{em,l}^a = R_{em} \left(\frac{T_{em}^2}{k^2} + \frac{2c_f}{k^2} \omega_{em} T_{em} + \frac{c_f^2}{k^2} \omega_{em}^2 \right) + c_f \omega_{em}^2 \quad (10)$$

This model is fitted to the measured data of the losses given used in the map based model in Section 2.1. The parameters in the analytical model are found by minimizing the least square error between (1) and (10), and the parameters k , R_{em} , and c_f are found to be 0.50 Nm/A , $0.065 \text{ }\Omega$, and 0.0029 Nm/s , respectively. The battery voltage is assumed to be the open circuit voltage, i.e. 250 V , when using the map to find the losses. The power losses computed in (10) are compared with the measured losses in Figure 1.

The power electronics is assumed to be an ideal component also in this model, and the expression for the battery current is given by

$$I_b = \frac{I_{em}U_{em}}{U_b} \quad (11)$$

CONTROLLER

A torque from the electric machine is requested from the energy management operating on vehicle level. The controller of the machine computes a requested

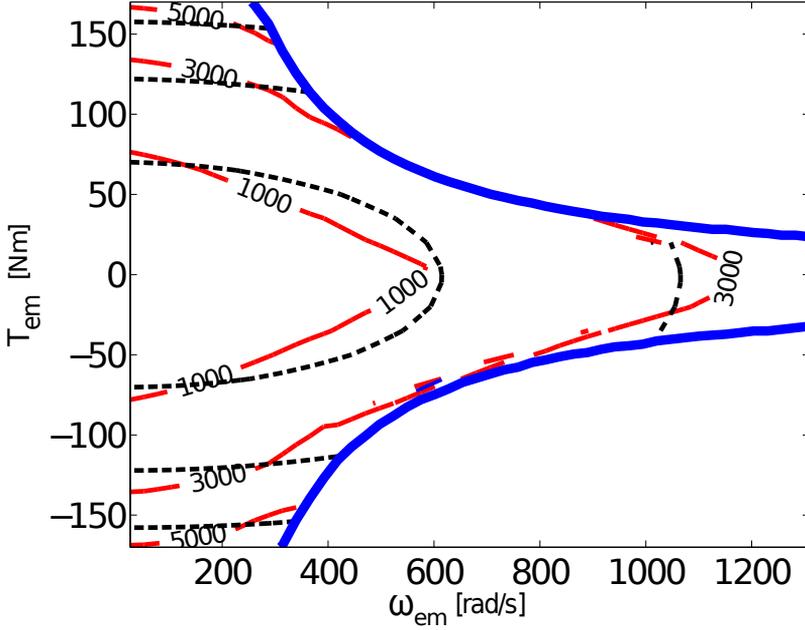


Figure 1: The power losses [W] of the machine and power electronics. The solid (thin) red lines show the measured losses in the map described in Section 2.1, the dashed lines the losses in the model described in Section 2.2, and the solid (thick) blue line the torque limitation of the machine.

voltage, U_{em}^{ctrl} , from the power electronics in order for the machine to, if possible, deliver this torque. The controller is an open loop controller and $U_{em,ctrl}$ is computed by

$$U_{em}^{ctrl} = \left(\frac{T_{em,req}}{k} + \frac{c_f}{k} \omega_{em} \right) R_{em} + k \omega_{em} \quad (12)$$

The model for the power electronics supplies this voltage to the machine when the component is fault free, i.e.

$$U_{em} = U_{em,ctrl} \quad (13)$$

3 COMBINING THE MAP AND ANALYTICAL MODELS FOR FAULT MODELING

As stated above, the map based model presented in Section 2.1 is beneficial to model the nominal behavior of the machine due to its high accuracy. However,

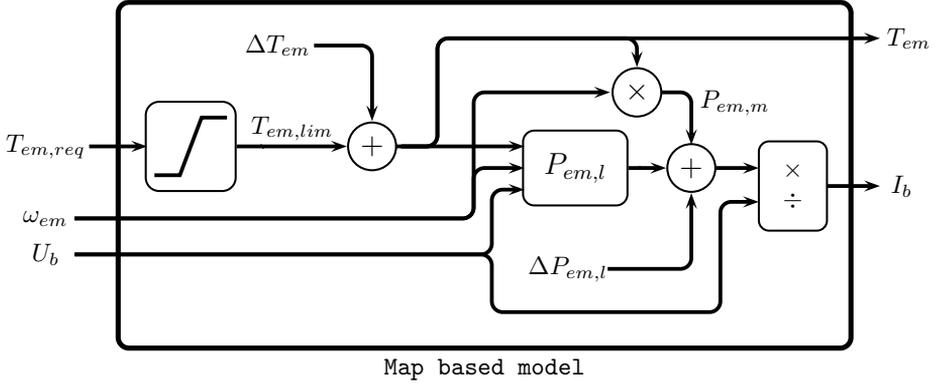


Figure 2: The map based model includes a limitation in the torque signal, since the machine has limitations in the torque it is capable to deliver, and the battery current is calculated from the mechanical power and the power losses. The map based model is extended with ΔT_{em} and $\Delta P_{em,l}$ to add the possibility to model faults in the machine. Note that the dynamics in the model in (3) is not included in the figure.

the model has the disadvantage that the parameters affected when a fault has occurred are not explicitly included in the model. In the fault free case, the map based model of the machine delivers the requested torque, as long as the machine is capable of delivering the torque, as can be seen in the schematic structure of the model in Figure 2. The battery current, I_b , is calculated using the mechanical power, $P_{em,m}$, and the power losses, $P_{em,l}$, that is a map and depends on the operating points of the machine, as described in Section 2.1.

The two main ways to model faults in a map based model is to modify the input or output signals of the map. The model is here extended to modify the delivered torque from the machine when a fault has occurred, by modifying the requested torque using ΔT_{em} according to Figure 2. This results in that the power losses of the machine changes when there is a fault affecting the delivered torque. A fault affecting the power losses of the machine affect the battery current, and is modeled using $\Delta P_{em,l}$. Expressions for ΔT_{em} and $\Delta P_{em,l}$ are derived in Sections 3.1 and 3.2 respectively.

It is three fault modes that are considered in the design of the diagnosis systems described in Section 5, and these faults are also used to evaluate the diagnosis system in simulations. Two of the faults affect the electric machine, by modifications in the resistance of the machine and the lumped torque and speed constant k used in the analytical model. A fault in the power electronics is modeled to result in that the applied voltage on the electric machine is not

the requested voltage. The faults are modeled as

$$k = k^{nom}(1 + f_{em,k}) \quad (14a)$$

$$R_{em} = R_{em}^{nom}(1 + f_{em,R}) \quad (14b)$$

$$U_{em} = U_{em}^{NF}(1 + f_{pe}) \quad (14c)$$

where k^{nom} and R_{em}^{nom} are the nominal values of the parameters, and U_{em}^{NF} the delivered voltage from the power electronics in the fault free case. These faults are important to monitor since they affect the delivered torque from the electric machine, as can be seen by combining (7) and (8) for the analytical model

$$T_{em}^a = k \left(\frac{U_{em}}{R_{em}} - \frac{k}{R_{em}} \omega_{em} \right) - c_f \omega_{em} \quad (15)$$

In the analytical model of the electric machine described in Section 2.2, it is straightforward to induce the faults described in (14) since these parameters are included in the model. The accuracy is however generally lower in this category of models compared to the map based model. Therefore, the map based model is used to model the fault free case, and the analytical model is used to model the influence of the faults on the electrical machine.

3.1 FINDING AN EXPRESSION FOR ΔT_{em}

From (15) it is seen that all three fault modes in (14) affect the delivered torque of the electric machine, and is here modeled according to

$$T_{em} = T_{em,lim} + \Delta T_{em} \quad (16)$$

where ΔT_{em} is the difference between T_{em}^{req} and T_{em} due to a fault in the system when $T_{em,req} = T_{em,lim}$. To find the expression for ΔT_{em} , the torque delivered by the faulty machine is computed using (15), and the delivered torque in the fault free case, $T_{em}^{a,NF}$, is also computed using (15), but with the nominal values of the parameters in the machine. The parameters k and R_{em} , and the voltage U_{em} used to calculate T_{em}^a , include models for the faults according to (14), and ΔT_{em} is expressed by

$$\begin{aligned} \Delta T_{em} &= T_{em}^a - T_{em}^{a,NF} \\ &= \frac{k}{R_{em}} (U_{em} - k\omega_{em}) - \frac{k^{nom}}{R_{em}^{nom}} (U_{em}^{NF} - k^{nom}\omega_{em}) \end{aligned} \quad (17)$$

The voltage U_{em}^{ctrl} needs to be calculated to find U_{em} and U_{em}^{NF} used in (17). This voltage is however not modeled in the map based model, and is therefore computed using the controller in analytical model given in (12). Information about if the system is faulty or fault free is not assumed to be known in the controller of the machine where U_{em}^{ctrl} is set, and therefore R_{em} and k in the expression are the nominal values even if there is a fault in the machine affecting

these parameters. The expression for U_{em}^{ctrl} used to compute ΔT_{em} and $\Delta P_{em,l}$ is thereby

$$U_{em}^{ctrl} = \left(\frac{T_{em,req}}{k^{nom}} + \frac{c_f}{k^{nom}} \omega_{em} \right) R_{em}^{nom} + k^{nom} \omega_{em} \quad (18)$$

3.2 FINDING AN EXPRESSION FOR $\Delta P_{em,l}$

The expression for the power losses in the analytical model is given in (10), and the expression states that $f_{em,k}$ and $f_{em,R}$ affect the power losses in the model. The losses in the map based model are modeled as

$$\tilde{P}_{em,l}^{map} = P_{em,l}^{map} + \Delta P_{em,l} \quad (19)$$

where $P_{em,l}^{map}$ is the original map and $\Delta P_{em,l}$ describes the difference in the power losses in the machine in the fault free and faulty cases of the machine. The losses in the faulty case are computed by (10), and the losses in the fault free case, $P_{em,l}^{a,NF}$, are also computed by (10), but with the nominal values of the parameters R_{em} and k . The torque used in the computations of $P_{em,l}^a$ and $P_{em,l}^{a,NF}$ is the delivered torque T_{em} from the machine in the map based model, see Figure 2. The modifications in the power losses is computed by

$$\begin{aligned} \Delta P_{em,l} &= P_{em,l}^a - P_{em,l}^{a,NF} \\ &= R_{em} \left(\frac{T_{em}^2}{k^2} + \frac{2c_f}{k^2} \omega_{em} T_{em} + \frac{c_f^2}{k^2} \omega_{em}^2 \right) + c_{f,em} \omega_{em}^2 - \\ &\quad - \left[R_{em}^{nom} \left(\frac{T_{em}^2}{(k^{nom})^2} + \frac{2c_f}{(k^{nom})^2} \omega_{em} T_{em} + \frac{c_f^2}{(k^{nom})^2} \omega_{em}^2 \right) + c_{f,em} \omega_{em}^2 \right] \\ &= \left(\frac{R_{em}}{k^2} - \frac{R_{em}^{nom}}{(k^{nom})^2} \right) (T_{em}^2 + 2c_f \omega_{em} T_{em} + c_f^2 \omega_{em}^2) \quad (20) \end{aligned}$$

4 ISOLABILITY GAIN BY COMBINING MODELS

In this section the maximum theoretical isolability performance of a diagnosis system based on the map based model is discussed. Firstly, the isolability performance using only the map based model is considered, i.e. the results from Section 3 are assumed not to be known. Secondly, the performance when combining the map based model with the fault models obtained in Section 3 are considered.

4.1 THEORETICAL FAULT ISOLABILITY USING MAP BASED MODEL

First, consider the case when using only the map based model, without any fault models. There are three fault modes to be monitored, see (14), and a single-fault

assumption is made in the diagnosis system design. On the basis of using only the map based model, it is reasonable that all three faults affect the delivered torque

$$T_{em} = g_1(f_{em,k}, f_{em,R}, f_{pe}) \quad (21)$$

while the power losses only depend on the fault modes in the electric machine

$$P_{em,l} = g_2(f_{em,k}, f_{em,R}) \quad (22)$$

Note that g_1 and g_2 also depend on other variables.

When equation (22) for $P_{em,l}$ is not consistent, this can be explained by either of the two faults $f_{em,k}$ or $f_{em,R}$. The fault f_{pe} can not be the cause since it does not affect the power losses and a single fault assumption is made. If the equation for T_{em} is inconsistent, this can be explained with any of the faults according to (21). Therefore a fault in the electric machine can be isolated from a fault in the power electronics, but not vice versa. Further, it is not possible to isolate the fault modes in the electric machine from each other when no fault models are used, since both these faults affect the same model equations.

4.2 THEORETICAL FAULT ISOLABILITY USING A COMBINED MODEL

Now consider the case where fault models are used in combination with the map. As stated in Section 4.1, fault models are required to isolate the fault modes from each other in the diagnosis system. Here the faults' influence on ΔT_{em} and $\Delta P_{em,l}$ described in (17) and (20) are used in the diagnosis system. It is assumed that the faults are constant or slowly varying, and is modeled as $\dot{f} = 0$. Note that the parameters k and R_{em} , and the voltage U_{em} all are included in the expression for T_{em} , and that the faults affect the torque in different ways. This means that full fault isolability can be achieved using only information about how T_{em} is modified i.e. by only using (14) and (17). The information from how the faults affect $\Delta P_{em,l}$ is however also used in the estimation of the faults using observers in the next section.

5 DESIGN OF A DIAGNOSIS SYSTEM

A diagnosis system monitoring the power electronics and the electric machine of the HEV is designed based on the models presented in Sections 2 and 3. The model used, including the fault models, is first transformed into state space form in Section 5.1. In Section 5.2, observers are designed based on the combined map and analytical model, but also based on only the analytical model. Based on the estimated faults from the observers, residual generators and test quantities are finally designed and evaluated in Section 5.3.

5.1 STATE SPACE FORMULATION OF THE MODEL

The model considered below is the combined map and analytical model of the electric machine. The model used in the design of the diagnosis systems is static since the dynamics in (3) is fast and is here assumed instantaneous. In the model used, the faults are included as states that are estimated in the observers. In its original form the model is a DAE of differential index one Petzold and Ascher (1998), but is reformulated as an ODE to be able to use standard observer techniques. The model is in the form

$$x_{1,t+1} = x_{1,t} + \omega_t \quad (23a)$$

$$0 = g(x_{1,t}, x_{2,t}, u_t) \quad (23b)$$

$$y_t = h(x_{1,t}, x_{2,t}, u_t) \quad (23c)$$

where x_1 is the vector of the dynamic variables, which is equal to the three faults, x_2 is the vector of algebraic variables, u is the vector of known input signals, and ω noise describing the model uncertainty. The expression $g(x_1, x_2, u)$ includes the model equations, and, since the DAE has differential index one, the algebraic variables x_2 can be computed from $g(x_1, x_2, u)$ by

$$x_{2,t} = g^{-1}(x_{1,t}, u_t) = G(x_{1,t}, u_t) \quad (24)$$

leading to the ODE

$$x_{1,t} = x_{1,t-1} + \omega_t \quad (25a)$$

$$y_t = h(x_{1,t}, G(x_{1,t}, u_t), u_t) \quad (25b)$$

which has the same solution set as (23). The algebraic variables x_2 and $G(x_1, u)$ are given by

$$\underbrace{\begin{bmatrix} k \\ R_{em} \\ T_{em,lim} \\ U_{em}^{ctrl} \\ U_{em} \\ U_{em}^{NF} \\ \Delta T_{em} \\ T_{em} \\ P_{em,m} \\ \Delta P_{em,l} \\ P_{em,l}^{map} \\ P_{em,e} \\ I_b \end{bmatrix}}_{x_2} = \underbrace{\begin{bmatrix} k^{nom}(1 + f_{em,k}) \\ R_{em}^{nom}(1 + f_{em,R}) \\ \min\{\max\{T_{em,min}, T_{em,req}\}, T_{em,max}\} \\ \frac{R_{em}^{nom}}{k^{nom}}(T_{em,lim} + c_f \omega_{em}) + k^{nom} \omega_{em} \\ U_{em}^{ctrl}(1 + f_{pe}) \\ U_{em}^{ctrl} \\ \frac{k}{R_{em}}[U_{em} - k \omega_{em}] - \frac{k^{nom}}{R_{em}^{nom}}[U_{em}^{NF} - k^{nom} \omega_{em}] \\ T_{em,lim} + \Delta T_{em} \\ T_{em} \omega_{em} \\ \left[\frac{R_{em}}{k^2} - \frac{R_{em}^{nom}}{(k^{nom})^2} \right] [T_{em}^2 + 2c_f \omega_{em} T_{em} + c_f^2 \omega_{em}^2] \\ f(T_{em}, \omega_{em}, U_b) \\ P_{em,m} + P_{em,l}^{map} + \Delta P_{em,l} \\ \frac{P_{em,e}}{U_b} \end{bmatrix}}_{G(x_1, u)} \quad (26)$$

The input signals are the requested torque, angular speed, and the battery voltage. Of these the angular speed and battery voltage are sensor signals

$$u = \begin{bmatrix} T_{em,req} \\ \omega_{em} \\ U_b \end{bmatrix} \quad (27)$$

The output signals are the delivered torque and the battery current, that are calculated in (23c) and are given by

$$h(x_1, x_2, u) = \begin{bmatrix} T_{em} \\ I_b \end{bmatrix} \quad (28)$$

5.2 FAULT ESTIMATION

The designed diagnosis systems are based on state-observers estimating the considered faults in (14). Several ways of doing this is performed and evaluated in this section, and the straightforward approach is to estimate all three faults in one observer.

ONE OBSERVER

The observer is an Extended Kalman Filter (EKF) Kailath et al. (2000) and the model equations are given by (26). The observer is denoted as \mathcal{O}_{all} , and is described by

$$\mathcal{O}_{all} : \begin{cases} \hat{x}_{1,t+1} = \hat{x}_{1,t} + K(y_t - \hat{y}_t) \\ y_t = h(\hat{x}_{1,t}, G(\hat{x}_{1,t}, u_t), u_t) \end{cases} \quad (29)$$

where $x_1 = [f_{em,k} \ f_{em,R} \ f_{pe}]^T$. These states are observable since the faults affect the output signals in different ways, see the model in (26) and (28).

The sensors used in the feedback term in the observer are a machine torque sensor and a battery current sensor. The sensor model includes Gaussian distributed white noise, $\nu_{T_{em}}$ and ν_{I_b} respectively, that is added to the signal the sensor measure

$$y = \begin{bmatrix} T_{em,sens} \\ I_{b,sens} \end{bmatrix} = \begin{bmatrix} T_{em} + \nu_{T_{em}} \\ I_b + \nu_{I_b} \end{bmatrix} \quad (30)$$

The torque sensor is used for simplicity, but if it is not available it is possible to use other sensors and extend the model used in the observers, see e.g. Sundström (2011), where also the observer design is given.

To evaluate the designed observer, the faults in (14) are induced one by one in the vehicle model and the observer estimates the faults. The faults are induced after 400 seconds in the simulation, and the sizes of the induced faults are $f_{em,k} = -0.03$, $f_{em,R} = -0.03$, and $f_{pe} = -0.01$. The driving cycle used is FTP75 and the estimated faults when $f_{em,k}$ has occurred, i.e. $x_1 = [-0.03 \ 0 \ 0]^T$, are presented in Figure 3. As can be seen in the figure, all faults are nonzero

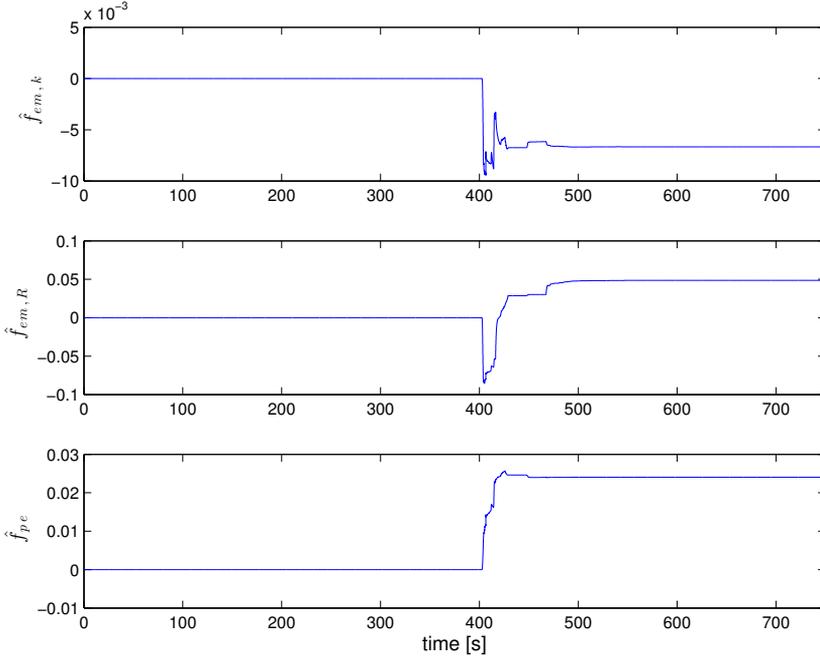


Figure 3: Estimated faults when $k = 0.97k_{nom}$ using \mathcal{O}_{all} .

after the fault is induced, and the estimated faults have converged to almost constant values that are $\hat{x}_1 = \{\hat{f}_{em,k}, \hat{f}_{em,R}, \hat{f}_{pe}\} = \{-0.0067, 0.049, 0.024\}$. Due to the erroneous estimation of the states in the observer presented in Figure 3, that appear even though the faults are observable as stated above, the difference in the output signals T_{em} and I_b based on \hat{x}_1 and the true fault values, $x_1^* = [-0.03 \ 0 \ 0]^T$ is investigated. The output signals computed based on \hat{x}_1 are denoted as \hat{T}_{em} and \hat{I}_b , and the output signals based on x_1^* are denoted as T_{em}^* and I_b^* . The difference between the outputs \hat{T}_{em} and T_{em}^* , and \hat{I}_b and I_b^* are computed as

$$\Delta_{T_{em}} = \hat{T}_{em} - T_{em}^* \quad (31a)$$

$$\Delta_{I_b} = \hat{I}_b - I_b^* \quad (31b)$$

and the differences are shown in Figure 4. Note that the maximum magnitudes of $\Delta_{T_{em}}$ and Δ_{I_b} are 0.002 Nm and 0.01 A respectively, which corresponds to as small values as 0.01% of the maximum magnitude of both output signals. Thereby it can be stated that it is difficult to estimate the correct value of the faults using this observer.

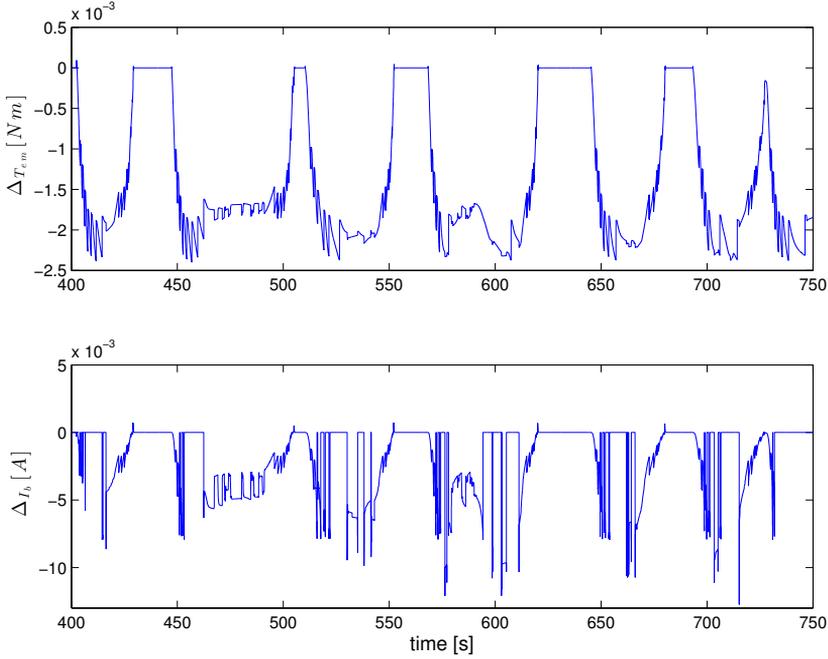


Figure 4: The differences ΔI_b and ΔT_{em} computed in (31). The difference in delivered torque and current is small, and therefore it is difficult to correctly estimate the faults.

THREE OBSERVERS ESTIMATING ONE FAULT EACH

Since the faults are almost not observable using one observer, the single-fault assumption is here used and three observers are designed to estimate one fault each. In each observer it is assumed that the other two faults are zero, and the three observers estimating $f_{em,k}$, $f_{em,R}$, and f_{pe} , are denoted $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} respectively. The basic idea to achieve fault detection and isolation based on these observers is described in Section 5.3, but first the observers are briefly described.

All three observers $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} use the same model equations, except for which fault that is to be estimated, and the used model equations are given by (26). As stated above, two faults in (26) are assumed to be zero in the observers, and x_1 in (25a) only includes the fault that is to be estimated in the observer. The observer used to estimate the fault $f_{em,k}$ is e.g.

$$\mathcal{O}_{f_{em,k}} : \begin{cases} \hat{f}_{em,k,t+1} = \hat{f}_{em,k,t} + K(y_t - \hat{y}_t) \\ y_t = h\left(\begin{bmatrix} \hat{f}_{em,k,t} \\ 0 \\ 0 \end{bmatrix}, G\left(\begin{bmatrix} \hat{f}_{em,k,t} \\ 0 \\ 0 \end{bmatrix}, u_t\right), u_t\right) \end{cases} \quad (32)$$

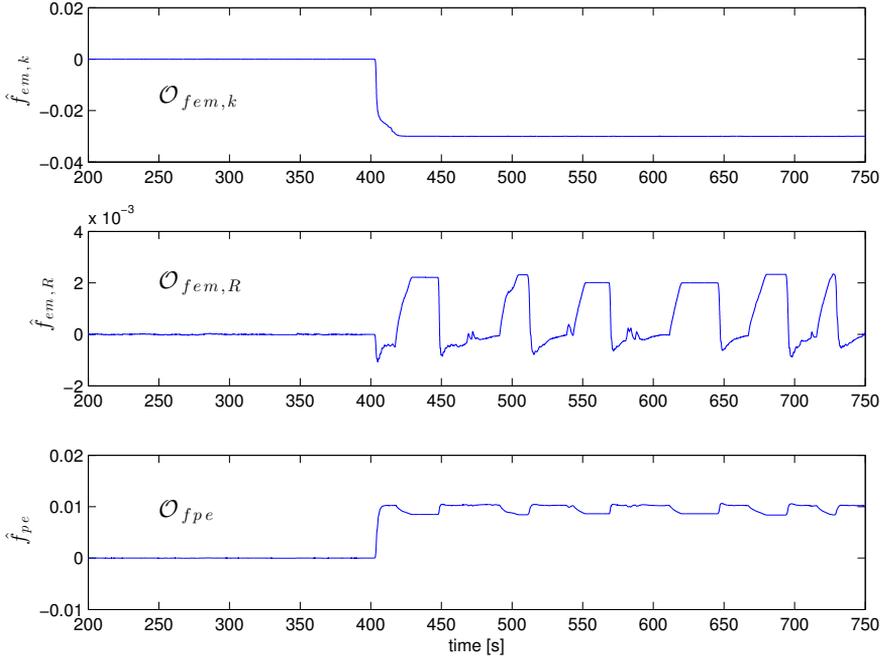


Figure 5: The estimated faults in the three observers $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} when $k = 0.97k^{nom}$.

since $\hat{x}_1 = [\hat{f}_{em,k} \ 0 \ 0]^T$. The sensors used are given by (30). Note that it is only one of the three estimated faults in the observers that ideally estimates a correct value of the fault. E.g. when $f_{em,k}$ has occurred, $\hat{f}_{em,R}$ and \hat{f}_{pe} estimated in $\mathcal{O}_{em,R}$ and \mathcal{O}_{pe} respectively, will possibly take any value even though the actual faults $f_{em,R} = f_{pe} = 0$.

The faults are induced one by one in the vehicle model to evaluate the designed observers. The estimated faults using the three observers when $\{f_{em,k}, f_{em,R}, f_{pe}\} = \{-0.03, 0, 0\}$ are given in Figure 5. In the upper plot, $\hat{f}_{em,k}$ is shown and it can be seen that the fault is accurately estimated using $\mathcal{O}_{em,k}$. The lower two plots present $\hat{f}_{em,R}$ and \hat{f}_{pe} , and since the induced fault in the simulation is $f_{em,k}$ and the three faults affect the model (26) differently, $\hat{f}_{em,R}$ and \hat{f}_{pe} do not converge to constant values. The reason is that $\hat{f}_{em,R}$ and \hat{f}_{pe} varies with the operating point of the machine to achieve the same outputs of the system as the measurements do. The information about what estimated faults that converge to constant values will later be used in the diagnosis system design to pinpoint what fault that has occurred. The results when inducing any of $f_{em,R}$ and f_{pe} in the vehicle model are similar to the results in Figure 5 when $f_{em,k}$ is induced.

THREE OBSERVERS BASED ON ONLY THE ANALYTICAL MODEL

In the previous two sections the combined map and analytical model, as described in Section 3, is used in the observers. In Figure 1 it is found that the analytical model does not accurately describe the power losses of the machine. How this modeling error affect the diagnosis performance is here investigated by designing three observers based on only the analytical model described in Section 2.2. The inputs u and outputs y are the same as in (27) and (28) respectively. The algebraic variables are denoted as x_2^a , and these are computed from a set of equations $G^a(x_1, u)$ as

$$\underbrace{\begin{bmatrix} k \\ R_{em} \\ U_{em}^{ctrl} \\ U_{em} \\ I_{em} \\ T_{em} \\ I_b \end{bmatrix}}_{x_2^a} = \underbrace{\begin{bmatrix} k^{nom}(1 + f_{em,k}) \\ R_{em}^{nom}(1 + f_{em,R}) \\ \frac{R_{em}^{nom}}{k^{nom}}(T_{em,req} + c_f \omega_{em}) + k^{nom} \omega_{em} \\ U_{em}^{ctrl}(1 + f_{pe}) \\ \frac{1}{R_{em}}(U_{em} - k \omega_{em}) \\ k I_{em} - c_f \omega_{em} \\ \frac{U_{em} I_{em}}{U_b} \end{bmatrix}}_{G^a(x_1, u)} \quad (33)$$

Three observers are designed estimating one fault each, and the observers are denoted as $\mathcal{O}_{em,k}^a$, $\mathcal{O}_{em,R}^a$, and \mathcal{O}_{pe}^a . The estimated faults when $f_{em,k} = -0.03$ are seen in Figure 6.

By comparing Figures 5 and 6 it is clear that the different models used in the observers results in different values of the estimated faults. Note especially that the fault that is induced, $f_{em,k}$, is more accurately estimated in $\mathcal{O}_{f_{em,k}}$ compared to $\mathcal{O}_{f_{em,k}}^a$, see the upper plots in Figures 5 and 6. The reason for that the estimated fault $\hat{f}_{em,k}$ is less accurately estimated in $\mathcal{O}_{f_{em,k}}$ compared to $\mathcal{O}_{f_{em,k}}^a$ when the fault $f_{em,k}$ is induced, is that the analytical model does not describe the model outputs T_{em} and I_b as good as the combined model.

5.3 DESIGN OF RESIDUAL GENERATORS AND TEST QUANTITIES

When an estimated fault is nonzero the system is assumed faulty. Based on the estimated faults in the observers presented above it is thereby easy to detect a fault in the monitored system since the estimated faults clearly becomes nonzero fast after the fault is induced in the vehicle. However, to pinpoint what fault that has occurred is more difficult since the faults can not be simultaneously accurately estimated using one observer, \mathcal{O}_{all} . Therefore, to be able to isolate the faults, residual generators based on the estimated faults in $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} are used in the diagnosis system. To increase the diagnosis performance, the residuals are post processed to form test quantities as described below.

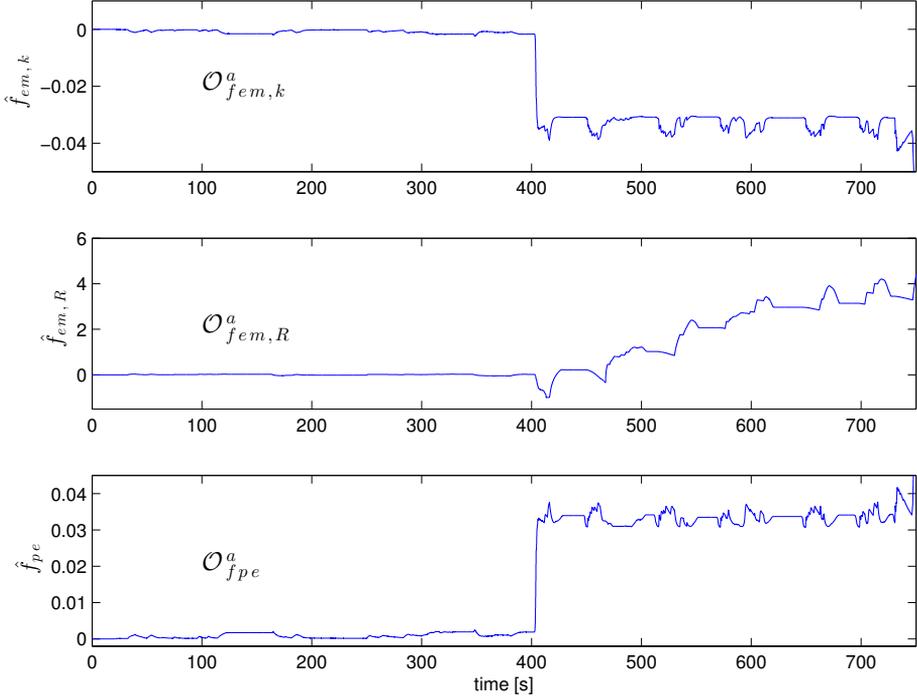


Figure 6: The estimated faults in the three observers based on the analytical model when $f_{em,k} = -0.03$.

RESIDUAL GENERATORS AND TEST QUANTITY FORMATION

The faults affect the system to be monitored in different ways, see (26). This, in combination with the assumption of constant faults, is used in the design of the residual generators and the basic idea is illustrated by an example.

Example 1. When a constant fault has occurred in the power electronics resulting in that $U_{em} \neq U_{em}^{ctrl}$, the estimate \hat{f}_{pe} is constant, but the estimated faults $\hat{f}_{em,k}$ and $\hat{f}_{em,R}$ calculated in $\mathcal{O}_{em,k}$ and $\mathcal{O}_{em,R}$, are dependent on the operating point of the electric machine. The reason for this is illustrated by an example using the expression for ΔT_{em} in (17). In this expression it is only T_{em}^a , and not $T_{em}^{a,NF}$, that is affected when there is a fault in the component. Combining (14) and (15) leads to

$$T_{em}^a = -c_f \omega_{em} + k^{nom}(1 + f_{em,k}) \cdot \left(\frac{U_{em}^{NF}(1 + f_{pe})}{R_{em}^{nom}(1 + f_{em,R})} - \frac{k^{nom}(1 + f_{em,k})}{R_{em}^{nom}(1 + f_{em,R})} \omega_{em} \right) \quad (34)$$

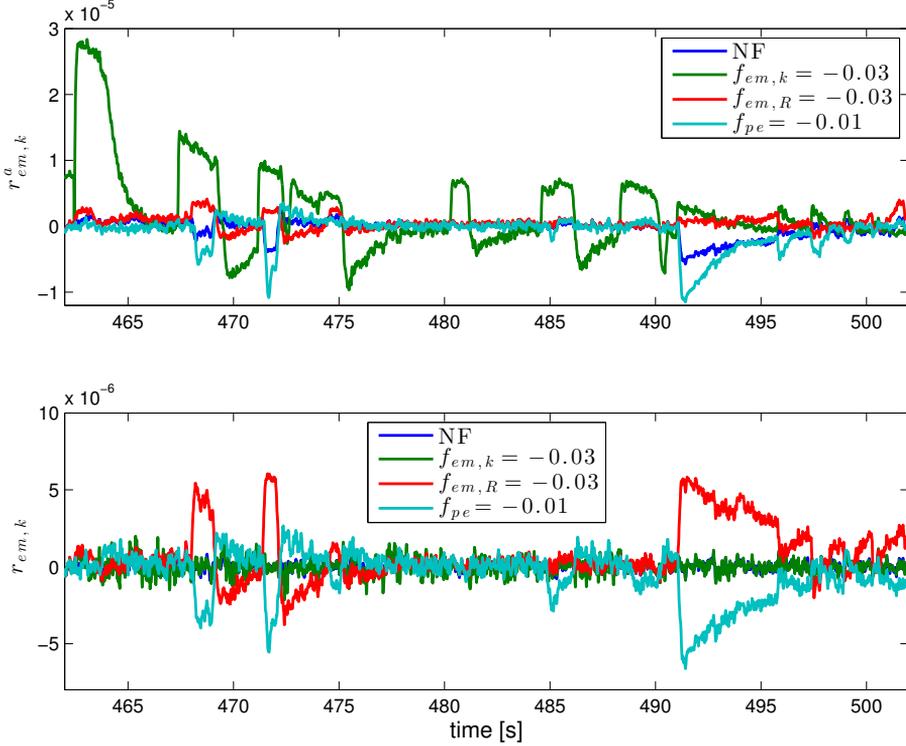


Figure 7: Residuals based on $\hat{f}_{em,k}$ from the observer based on only the analytical model, $\mathcal{O}_{em,k}^a$, in the upper plot and the observer based on the combined model, $\mathcal{O}_{em,k}$, in the lower plot. The residuals are computed both for the non-faulty case and when one fault at the time is induced in the vehicle model.

A fault affecting the resistance is e.g., included in two terms in the expression, one that is proportional to U_{em}^{NF} and one that is proportional to ω_{em} . The fault in the power electronics is only included in the term that is proportional to the voltage. This leads to that when there is a constant fault in the power electronics, the value of $\hat{f}_{em,R}$ varies with ω_{em} to achieve the same value for T_{em}^a as f_{pe} does. This information is used to construct residual generators in the diagnosis system. \square

The residual generators are designed based on that if an observer estimates a fault f_i , \hat{f}_i converges to f_i when this fault occurs, and that $\hat{f}_{i,t} = \hat{f}_{i,t-1}$ since f_i is assumed to be constant, i.e. $f_{i,t} = f_{i,t-1}$. However, if a fault mode $f_j \neq f_i$ occurs it is not possible to state anything about the value \hat{f}_i will take. Therefore, if $\hat{f}_{i,t} \neq \hat{f}_{i,t-1}$, this can only be explained with that there is a fault in the system and that this fault is not f_i . This idea is used in the design of three residual

Table 1: Decision structure for the diagnosis system including fault models. Full fault isolability is structurally achieved, since a unique set of tests ideally react for each fault.

	$f_{em,k}$	$f_{em,R}$	f_{pe}
$T1$		X	X
$T2$	X		X
$T3$	X	X	

generators that are based on the change in the estimated faults in the observers between two time steps

$$r_t = \hat{f}_{i,t} - \hat{f}_{i,t-1} \quad (35)$$

The residual generator based on $\hat{f}_{em,k}$ estimated in $\mathcal{O}_{em,k}$ is denoted as $r_{em,k}$, and this signal is presented in the lower plot in Figure 7 for different faults induced in the vehicle. As can be seen in the figure, $r_{em,k}$ reacts in a similar way in the fault free case and when $f_{em,k} = -0.03$. When any of the faults $f_{em,R}$ and f_{pe} occurs the signal is clearly separated from the fault free case, and therefore the residual has reacted, and this is as expected.

In Section 5.2 it is stated that the faults are more accurately estimated using the observers $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} , compared to $\mathcal{O}_{em,k}^a$, $\mathcal{O}_{em,R}^a$, and \mathcal{O}_{pe}^a , see Figures 5 and 6. To evaluate the influence of this in the diagnosis system, the corresponding residual to $r_{em,k}$ based on $\mathcal{O}_{em,k}^a$ is presented in the upper plot in Figure 7, and is denoted as $r_{em,k}^a$. This residual is not expected to react to the fault $f_{em,k}$, but as can be seen in the figure, the residual when $f_{em,k} = -0.03$ is significantly separated from the fault free case. To achieve a correct diagnostic decision there is a need to investigate what causes this separation, and to compensate for the difference in the residual. The results are the same for all three residuals based on the analytical model, and therefore it is verified that it is more time consuming to design a diagnosis system based on $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} , compared to $\mathcal{O}_{em,k}^a$, $\mathcal{O}_{em,R}^a$, and \mathcal{O}_{pe}^a . Due to this result, the residuals used in the further diagnosis system design process are based on the combined map and analytical models.

To reduce the impact of noise in the residuals in the decision making, the residuals are post processed to form test quantities. This is here done using the CUSUM algorithm Page (1954); Gustafsson (2000)

$$T_t = \max \{0, T_{t-1} + |r_t| - \nu\} \quad (36)$$

where ν is a design parameter that corresponds to the noise and model uncertainty in the residuals. A test reacts when T is above a threshold J . To evaluate the performance of a diagnosis system, normalized test quantities, T_{norm} , are calculated based on the maximum value, $T_{\text{max,NF}}$, of T in the fault free case. An alarm is generated when $T_{\text{norm}} > 1$ and the design parameter $\Phi \in [0, 1]$ states

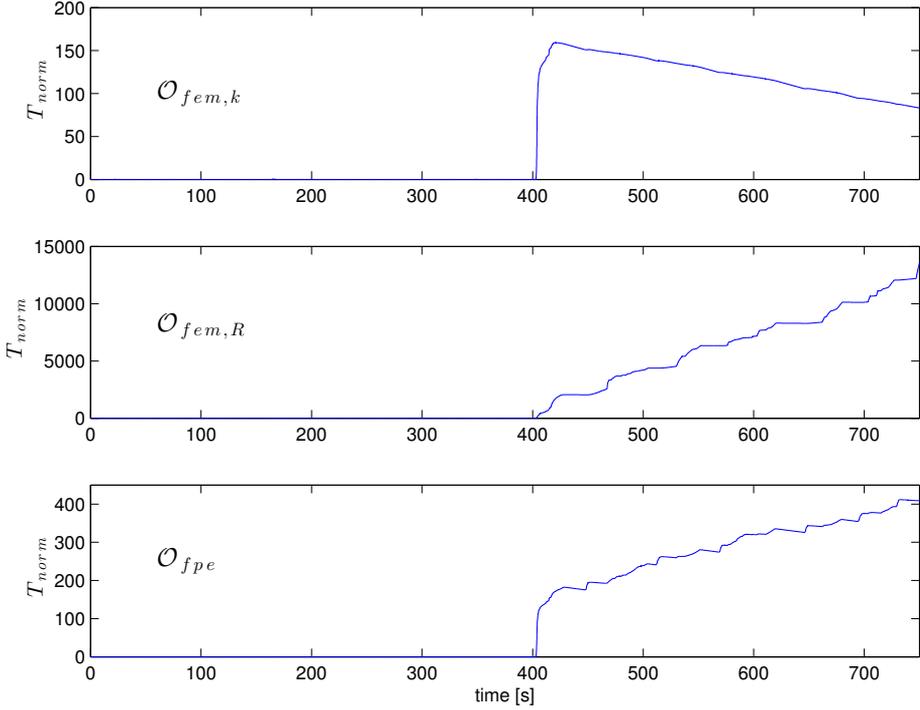


Figure 8: Test quantities based on the estimated faults presented in Figure 5 when $f_{em,k} = -0.03$.

the margin to false alarm, and is here set to 0.5

$$T_{\text{norm}} = \frac{T}{T_{\text{max,NF}}} \Phi \quad (37)$$

The decision structure includes information about which faults the different tests are expected to react to and is shown in Table 1. If e.g. T_1 has reacted this can be explained by $f_{em,R}$ or f_{pe} , but if also T_2 has reacted the only single fault diagnosis is a fault in the power electronics. The tests T_1 , T_2 , and T_3 are based on $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} respectively.

The normalized test quantities based on $\mathcal{O}_{em,k}$, $\mathcal{O}_{em,R}$, and \mathcal{O}_{pe} when $f_{em,k}$ is induced as a step in the vehicle model after 400 seconds are presented in Figure 8. As can be seen in the upper plot, $T_{\text{norm}} > 1$ when the fault is induced at 400 seconds. When a fault occurs in the system, the assumption $f_{i,t} = f_{i,t-1}$ is not valid. This may lead to that a residual that is not expected to react on the fault according to Table 1 becomes non-zero, and the corresponding test quantity reacts. To avoid this behavior it is needed to identify the time interval

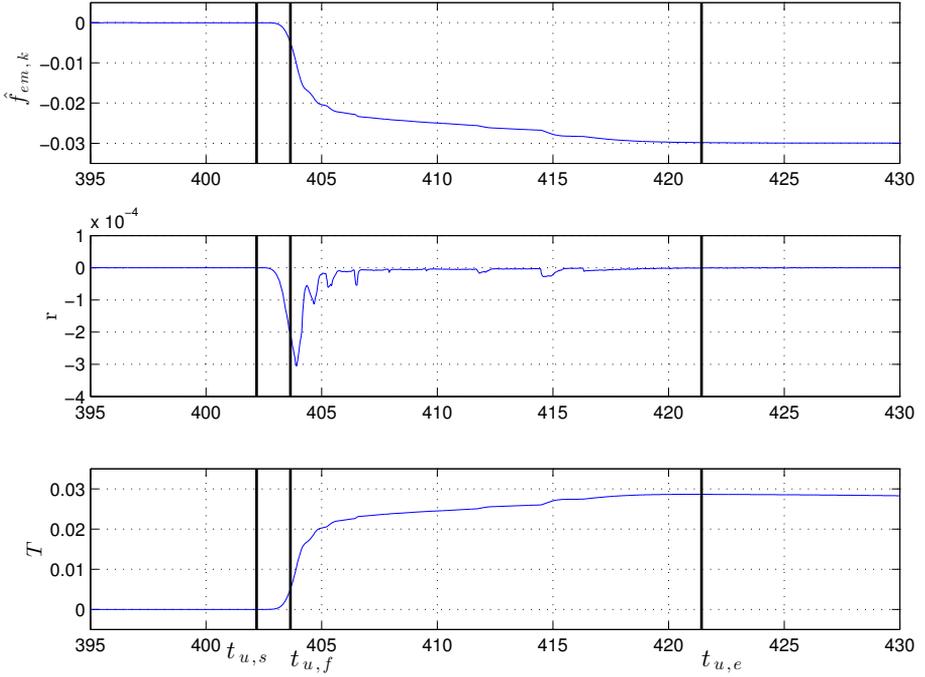


Figure 9: Illustration of the times used to state when to not update the test quantity exemplified with $\mathcal{O}_{f_{em,k}}$ when $f_{em,k} = -0.03$.

after a fault has occurred till the estimated state in the observer estimating the fault that has occurred has converged, in this case $\hat{f}_{em,k}$, to the new, faulty, value. During this time period the test quantity is not to be updated.

NOT UPDATE TEST QUANTITY DURING TRANSIENTS

The start time of the time period where the test quantities are not updated is denoted as $t_{u,s}$, the end time is denoted as $t_{u,e}$, and the time when the test quantity should not be updated is defined by $\tau_u = \{t_{u,s}, t_{u,e}\}$. Define all times in the driving cycle where the considered residual computed as in (35) changes sign as τ_r . To find the times $t_{u,s}$ and $t_{u,e}$, the first time the magnitude of the estimated fault is above a threshold J_f is used

$$t_{u,f} = \min\{t : |\hat{f}_{i,t}| > J_f\} \quad (38a)$$

It is wanted to start to update the test quantity when the estimated fault in the observer has converged to a the correct value of the fault. The residual is either positive or negative before \hat{f}_i has converged to the actual value of the fault, and

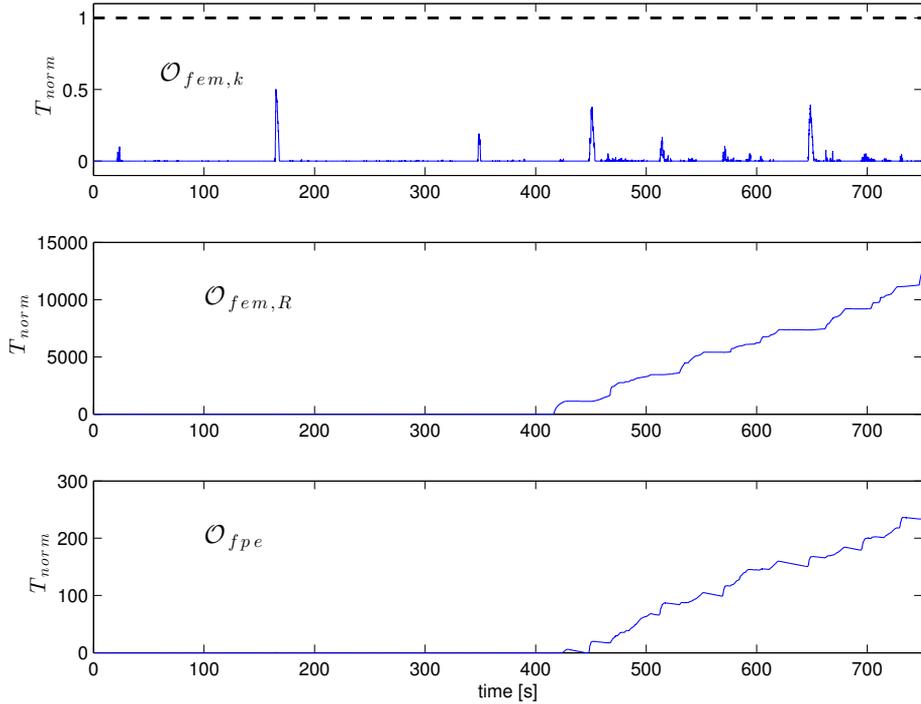


Figure 10: The test quantities based on the estimated faults in Figure 5, but the test quantities are not updated when the fault is induced in the model.

therefore the times $t_{u,s}$ and $t_{u,e}$ are computed by

$$t_{u,s} = \max\{\tau_r < t_{u,f}\} \quad (38b)$$

$$t_{u,e} = \min\{\tau_r > t_{u,f}\} \quad (38c)$$

In Figure 9 the estimated fault $\hat{f}_{em,k}$ is shown, as well as the corresponding residual and test quantities. In the figure the times $t_{u,f}$, $t_{u,s}$, and $t_{u,e}$ are also shown.

Figure 10 shows the test quantities when these are not updated according to (38). As can be seen in the upper plot there is no false alarm when the fault is induced in the system, and the other two tests react on the fault as expected.

5.4 SUMMING UP

In the case study the value of using a combination of two models, one that accurately describes the nominal behavior of the machine and one that describes how the faults affect the outputs, in a diagnosis system is demonstrated. Further,

a scheme of using several observers estimating one fault each is used to isolate the faults, and the scheme is based on the single-fault assumption and that $\dot{f} = 0$. At the time instant when a fault occurs in the system the fault is obviously not constant, and an approach to avoid false alarm due to this by not updating the test quantities is used, compare Figures 8 and 10. The method is successful even though the faults are induced as steps and not slowly increasing, which would be easier to handle since this is closer to the assumption $\dot{f} = 0$.

6 CONCLUSIONS

A method combining two models, one with good fault free accuracy and one that models how the faults affect the system, is used in the design of a diagnosis system for the power electronics and the electric machine in an HEV. An extensive simulation study shows that the whole approach works well including the specific scheme of using several observers estimating one fault each. Thus, a promising path to achieve both good fault detectability and isolability performance is presented, without the need for neither measurements of a faulty system nor detailed physical modeling.

REFERENCES

- M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer, 2nd edition, 2006.
- Z. Chen, Y. Fu, and C. Mi. State of charge estimation of lithium-ion batteries in electric drive vehicles using extended kalman filtering. *Vehicular Technology, IEEE Transactions on*, 62(3):1020–1030, 2013.
- J. de Kleer, A. Mackworth, and R. Reiter. Characterizing diagnoses and systems. *Artificial Intelligence*, 56(2-3):197–222, 1992.
- F. Gustafsson. *Adaptive filtering and change detection*. John Wiley & Sons, 2000.
- L. Guzzella and A. Sciarretta. *Vehicle Propulsion System, Introduction to Modeling and Optimization*. Springer Verlag, Zürich, 2 edition, 2007.
- T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice Hall, 2000.
- P. Mellor. High efficiency drive-trains for electric and hybrid vehicles. In *Electrical Machine Design for All-Electric and Hybrid-Electric Vehicles (Ref. No. 1999/196)*, *IEE Colloquium on*, pages 8/1 –8/7, 1999.
- E. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.
- L. R. Petzold and U. M. Ascher. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Siam, 1998.
- C. Sundström. Vehicle level diagnosis for hybrid powertrains. Technical report, 2011. Licentiate thesis LiU-TEK-LIC-2011:27, Thesis No. 1488.
- C. Sundström, E. Frisk, and L. Nielsen. Fault monitoring of the electric machine in a hybrid vehicle. In *7th IFAC Symposium on Advances in Automotive Control*, pages 548–553, Tokyo, Japan, 2013.
- G. Zhu, L.-A. Dessaint, O. Akhrif, and A. Kaddouri. Speed tracking control of a permanent-magnet synchronous motor with state and load torque observer. *Industrial Electronics, IEEE Transactions on*, 47(2):346 –355, 2000.
- Z. Zhu and D. Howe. Electrical machines and drives for electric, hybrid, and fuel cell vehicles. *Proceedings of the IEEE*, 95(4):746 –765, April 2007.

Linköping studies in science and technology, Dissertations
Division of Vehicular Systems
Department of Electrical Engineering
Linköping University

- No 1** Magnus Pettersson, *Driveline Modeling and Control*, 1997.
- No 2** Lars Eriksson, *Spark Advance Modeling and Control*, 1999.
- No 3** Mattias Nyberg, *Model Based Fault Diagnosis: Methods, Theory, and Automotive Engine Applications*, 1999.
- No 4** Erik Frisk, *Residual Generation for Fault Diagnosis*, 2001.
- No 5** Per Andersson, *Air Charge Estimation in Turbocharged Spark Ignition Engines*, 2005.
- No 6** Mattias Krysanter, *Design and Analysis of Diagnosis Systems Using Structural Methods*, 2006.
- No 7** Jonas Biteus, *Fault Isolation in Distributed Embedded Systems*, 2007.
- No 8** Ylva Nilsson, *Modelling for Fuel Optimal Control of a Variable Compression Engine*, 2007.
- No 9** Markus Klein, *Single-Zone Cylinder Pressure Modeling and Estimation for Heat Release Analysis of SI Engines*, 2007.
- No 10** Anders Fröberg, *Efficient Simulation and Optimal Control for Vehicle Propulsion*, 2008.
- No 11** Per Öberg, *A DAE Formulation for Multi-Zone Thermodynamic Models and its Application to CVCP Engines*, 2009.
- No 12** Johan Wahlström, *Control of EGR and VGT for Emission Control and Pumping Work Minimization in Diesel Engines*, 2009.
- No 13** Anna Pernestål, *Probabilistic Fault Diagnosis with Automotive Applications*, 2009.
- No 14** Erik Hellström, *Look-ahead Control of Heavy Vehicles*, 2010.
- No 15** Erik Höckerdal, *Model Error Compensation in ODE and DAE Estimators with Automotive Engine Applications*, 2011.

No 16 Carl Svärd, *Methods for Automated Design of Fault Detection and Isolation Systems with Automotive Applications*, 2012.

No 17 Oskar Leufvén, *Modeling for control of centrifugal compressors*, 2013.