TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Jarmo Nurmi

On Increasing the Automation Level of Heavy-Duty Hydraulic Manipulators with Condition Monitoring of the Hydraulic System and Energy-Optimised Redundancy Resolution

Julkaisu 1481 • Publication 1481

Tampere 2017

Tampereen teknillinen yliopisto. Julkaisu 1481
Tampere University of Technology. Publication 1481

Jarmo Nurmi

On Increasing the Automation Level of Heavy-Duty Hydraulic Manipulators with Condition Monitoring of the Hydraulic System and Energy-Optimised Redundancy Resolution

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Konetalo Building, Auditorium K1702, at Tampere University of Technology, on the 30th of June 2017, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2017

Doctoral candidate: Jarmo Nurmi
Laboratory of Automation and Hydraulic Engineering
Faculty of Engineering Sciences
Tampere University of Technology
Finland

Supervisor: Prof. Jouni Mattila
Laboratory of Automation and Hydraulic Engineering
Faculty of Engineering Sciences
Tampere University of Technology
Finland

Pre-examiners: Prof. Matti Pietola
School of Engineering, Department of Mechanical Engineering
Aalto University
Finland

Associate Prof. Huapeng Wu
School of Energy Systems, Department of Mechanical Engineering
Lappeenranta University of Technology
Finland

Opponents: Prof. Matti Pietola
School of Engineering, Department of Mechanical Engineering
Aalto University
Finland

Dr. Arto Peltomaa
Program Manager at DIMECC Oy
Finland

ISBN 978-952-15-3974-9 (printed)
ISBN 978-952-15-3976-3 (PDF)
ISSN 1459-2045

Abstract

Hydraulic manipulators on mobile machines are predominantly used for excavation and lifting applications at construction sites and for heavy-duty material handling in the forest industry due to their superior power-density and rugged nature. These manipulators are conventionally open-loop controlled by human operators who are sufficiently skilled to operate the machines. However, in the footsteps of pioneering original equipment manufacturers (OEMs) and to keep up with the intensifying demand for innovation, more and more mobile machine OEMs have a major interest in significantly increasing the automation level of their hydraulic manipulators and improving the operation of manipulators. In this thesis, robotic software-based functionalities in the form of model-based condition monitoring and energy-optimal redundancy resolution which facilitate increased automation level of hydraulic manipulators are proposed.

A condition monitoring system generally consists of software modules and sensors which co-operate harmonically and monitor the hydraulic system's health in real-time based on an indirect measure of this system's health. The premise is that when this condition monitoring system recognises that the system's health has deteriorated past a given threshold (in other words, when a minor fault is detected, such as a slowly increasing internal leakage of the hydraulic cylinder), the condition monitoring module issues an alarm to warn the system operator of the malfunction, and the module could ideally diagnose the fault cause. In addition, when faced with severe faults, such as an external leakage or an abruptly increasing internal leakage in the hydraulic system, an alarm from the condition monitoring system ensures that the machine is quickly halted to prevent any further damage to the machine or its surroundings.

The basic requirement in the design of such a condition monitoring system is to make sure that this system is robust and fault-sensitive. These properties are difficult to achieve in complex mobile hydraulic systems on hydraulic manipulators due to the modelling uncertainties affecting these systems. The modelling uncertainties affecting mobile hydraulic systems are specific compared with many other types of systems and are large because of the hydraulic system complexities, nonlinearities, discontinuities and inherently time-varying parameters. A feasible solution to this modelling uncertainty problem would be to either attenuate the effect of modelling errors on the performance of model-based condition monitoring or to develop improved non-model-based methods with increased fault-sensitivity. In this research work, the former model-based approach is taken. Adaptation of the model residual thresholds based on system operating points and reliable, load-independent system models are proposed as integral parts of the condition monitoring solution to the modelling uncertainty problem. These proposed solutions make the realisation of condition monitoring solutions more difficult on heavy-duty hydraulic manipulators compared with fixed-load manipulators, for example. These solutions are

covered in detail in a subset of the research publications appended to this thesis.

There is wide-spread interest from hydraulic manipulator OEMs in increasing the automation level of their hydraulic manipulators. Most often, this interest is related to semi-automation of repetitive work cycles to improve work productivity and operator workload circumstances. This robotic semi-automated approach involves resolving the kinematic redundancy of hydraulic manipulators to obtain motion references for the joint controller to enable desirable closed-loop controlled motions. Because conventional redundancy resolutions are usually sub-optimal at the hydraulic system level, a hydraulic energy-optimised, global redundancy resolution is proposed in this thesis for the first time. Kinematic redundancy is resolved energy optimally from the standpoint of the hydraulic system along a prescribed path for a typical 3-degrees-of-freedom (3-DOF) and 4-DOF hydraulic manipulator. Joint motions are also constrained based on the actuators' position, velocity and acceleration bounds in hydraulic manipulators in the proposed solution. This kinematic redundancy resolution topic is discussed in the last two research papers. Overall, both designed manipulator features, condition monitoring and energy-optimised redundancy resolution, are believed to be essential for increasing the automation of hydraulic manipulators.

Preface

This study was carried out at Tampere University of Technology (TUT) at the Laboratory of Automation and Hydraulic Engineering (AUT), formerly the Department of Intelligent Hydraulics and Automation (IHA), during the years 2011 – 2017. My sincere thanks go to my supervisor, Professor Jouni Mattila, for his encouraging guidance and for providing the opportunity to carry out this interesting dissertation work in a professional and friendly environment. I am also grateful for the facilities and equipment provided by the Laboratory of AUT and Professor Kalevi Huhtala, the Head of AUT. The knowledgeable help from the laboratory personnel and colleagues in setting up the experimental systems is likewise appreciated.

Support from colleagues at the department is highly acknowledged. I would like to thank, in particular, fellow colleagues M.Sc. Tuomo Kivelä, Dr. Janne Koivumäki, Dr. Janne Honkakorpi, MSc. Janne Tuominen from IHA and Dr. Juho Vihonen from the Department of Signal Processing at TUT, who have probably helped me the most during the course of this work and with whom I have had the great pleasure of working with on projects related to this dissertation. I am also grateful to all my other colleagues who have provided their invaluable advice and help during this work.

I am grateful to preliminary examiners Prof. Matti Pietola and Associate Prof. Huapeng Wu for their evaluation of this dissertation and their professional feedback. It is always great to receive constructive feedback on one's own work from prestigious professionals working in the field. I am also grateful to Prof. Pietola and Dr. Arto Peltomaa for agreeing to be my opponents.

The funding received from the doctoral program of Concurrent Mechanical Engineering of Finland and the Academy of Finland is greatly appreciated. Without it this work could not have been done.

I am also in debt to my family, close relatives and friends for the support they have given me during these years.

In Tampere,
June 2017

Jarmo Nurmi

Contents

Acronyms

CP	constant-pressure
DH	Denavit-Hartenberg
DOF	degrees-of-freedom
DP	dynamic programming
EKF	extended Kalman filter
FD	fault diagnosis
FDD	fault detection and diagnosis
FDI	fault detection and isolation
GSA	global sensitivity analysis
KF	Kalman filter
LS	load-sensing
MEMS	micro-electromechanical systems
NLP	nonlinear programming problem
OEM	original equipment manufacturer
RP	research problem
UIO	unknown input observer
UKF	unscented Kalman filter

List of Publications

P-I: Jarmo Nurmi, Jouni Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 1: Global sensitivity analysis," *International Journal of Fluid Power*, vol. 12, no. 3, pp. 41–51, November 2011, http://dx.doi.org/10.1080/14399776.2011.10781036.

P-II: Jarmo Nurmi, Jouni Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading condition, Part 2: Fault detection and isolation scheme," *International Journal of Fluid Power*, vol. 13, no. 1, pp. 17–27, March 2012, http://dx.doi.org/10.1080/14399776.2012.10781043.

P-III: Jarmo Nurmi, Janne Honkakorpi, Juho Vihonen, Jouni Mattila, "Micro-electromechanical system sensors in unscented Kalman filter-based condition monitoring of hydraulic systems," *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, Wollongong, Australia, July 9–12, 2013, pp. 354–361, http://dx.doi.org/10.1109/AIM.2013.6584117.

P-IV: Jarmo Nurmi, Jouni Mattila, "Detection and isolation of faults in mobile hydraulic valves based on a reduced-order model and adaptive thresholds," *Proceedings of the ASME/BATH Symposium on Fluid Power & Motion Control (FPMC)*, Sarasota, Florida, USA, October 6–9, 2013, http://dx.doi.org/10.1115/FPMC2013-4435.

P-V: Jarmo Nurmi, Jouni Mattila, "Global energy-optimised redundancy resolution in hydraulic manipulators using dynamic programming," *Automation in Construction*, vol. 73, pp. 120–134, January 2017, http://dx.doi.org/10.1016/j.autcon.2016.09.006.

P-VI: Jarmo Nurmi, Jouni Mattila, "Global energy-optimal redundancy resolution of hydraulic manipulators: Experimental results for a forestry manipulator." *Energies*, vol. 10, no. 5, 31 p., May 2017, http://dx.doi.org/10.3390/en10050647.

1 Introduction

Hydraulic manipulators are used in numerous applications in the forestry, agriculture, mining and construction industries for heavy-duty material handling (Figure 1.1). Some of the most well-known applications are in agricultural tractors, forest harvesters and forwarders, construction excavators and front-end loaders, amongst other heavy equipment. The ruggedness of hydraulics combined with the superior power density of hydraulic systems compared to other existing system technology render these machines particularly suitable for their purpose.

The serial-link hydraulic manipulators used in these applications are mechanically complicated rigid-body systems. They are complemented with complex hydraulic circuits which typically conceal numerous highly-elaborate components, including hydraulic valves with complex hydro-mechanical built-in functionalities, multiple hydraulic actuators, hydro-mechanical feedback systems, fluid filters and (variable displacement-controlled) hydraulic pumps which power the fluid circuitry through the prime mover, usually a diesel engine. These mechanical and hydraulic system aspects of the manipulators are often coupled through an electronic control system, that binds the systems together as one complex electro-hydro-mechanical system.

There is considerable interest in industry and academia in the increasing automation of these complex hydraulic manipulators. Automation in the traditional sense implies autonomous robotic control of the manipulators, but there is more to automation than that. With automation, there is a growing need for different software-based robotic functionalities, ranging from motion planning algorithms which generate proper motion references for automated operations and (particularly further in the advance towards autonomous machines), to condition monitoring algorithms which monitor the underlying complex hydraulic systems for faults, and which, in future, could prevent system failures from occurring.

In accordance with the interest in automation, as expected, hydraulic manipulators are already undergoing a technological transformation from a feature design perspective. In addition, novel software-based manipulator features (such as coordinated manipulator tip control [2–4] for improved productivity and longer manipulator wear-life), as well as conventional structural improvements (such as better positioning of the manipulator [5] for greater outside visibility), have recently been developed. Improved motion sensors have also been introduced to hydraulic manipulator applications. This development is highly encouraging as it thrusts the industry towards autonomous robotic control applications.

The demand for product innovation in hydraulic manipulators is strong because original equipment manufacturers (OEMs) need additional standout features to differentiate themselves in the global market sector. For this reason, OEMs are particularly investigating

Figure 1.1: Heavy-duty hydraulic manipulator on a truck [1].

and shifting towards designing more of these automated products, where improvements can be had even without major system modifications through a properly orchestrated interplay between hardware sensors and software components.

1.1 Research Problem

There is high academic and industrial interest in robotic control of hydraulic manipulators and accordingly a growing need for robotic software-based functionalities which can improve the manipulator's operation through increasing automation. This interest encourages and motivates scholars to pursuit research in this area. The main source of motivation for conducting research in this area is that research on robotic software-based functionalities is lacking in the literature. We examine two major underdeveloped robotic functionalities of hydraulic manipulators: model-based condition monitoring of hydraulic components and redundancy resolution for robotic motion planning of hydraulic manipulators. To design these robotic functionalities for heavy-duty hydraulic manipulators, in particular, we focus on manipulators' typical valve-controlled hydraulic systems which are complex with characteristically high time-variant loads on the hydraulic cylinders and in which system energy inefficiency is a major problem [6].

Hence, as a partial solution to the research problem (RP), condition monitoring solutions are designed for hydraulic heavy-duty manipulators with the following features:

RP1: **Load-independency**: The condition monitoring solution should be load-independent because the valve-controlled hydraulic cylinders on heavy-duty manipulators are usually subjected to varying load conditions. Thus, the

solution should not require knowledge about the magnitude and variance of the payload disturbance at the manipulator end-effector, which is observed at the hydraulic cylinders.

RP2: **General applicability**: The condition monitoring solution should be applicable to multi-actuator hydraulic manipulators which are valve-controlled using proportional hydraulic valves or mobile hydraulic valves with complex built-in hydro-mechanical functionalities.

RP3: **High performance**: The condition monitoring solution should be fault-sensitive and capable of isolating common hydraulic system faults. The FDI performance should be experimentally evaluated.

As another partial solution to the RP, kinematic redundancy resolution is proposed for hydraulic heavy-duty manipulators with the following features:

RP4: **Energy-optimality**: Kinematic redundancy resolution should be energy-optimal at the valve-controlled hydraulic actuator and hydraulic power system interaction level. The minimum attainable energy consumption of the commonly-used load-sensing (LS) and constant-pressure (CP) hydraulic systems should be obtained.

RP5: **General applicability**: Kinematic redundancy resolution should be applicable to typical kinematically-redundant three-degrees-of-freedom (3-DOF) and 4-DOF hydraulic manipulator designs.

RP6: **High performance**: Kinematic redundancy resolution should provide high performance whilst satisfying the position, velocity and acceleration bounds of the actuators. The performance of the redundancy resolution should be experimentally assessed by comparison to conventional methods.

To summarise the main research problem, load-independent model-based condition monitoring solutions and energy-optimised redundancy resolution should be designed for hydraulic manipulators to improve the operation of these manipulators towards the realisation of robotic manipulator control. These solutions are generally not available on heavy-duty hydraulic manipulators predominantly operated by humans and are lacking in the literature. With the increased complexity and efficiency requirements affecting these manipulators and the fast-paced maturation of high-technology, low-cost, robust and even wireless embedded sensors, robotic control utilising these features is becoming a feasible reality.

Condition monitoring and redundancy resolution are instrumental robotic functionalities which increase the automation of hydraulic manipulators. Why these robotic functionalities, in particular, are important enablers of increased automation is presented in more detail in the following sections.

1.2 Motivation for Condition Monitoring of Hydraulic Components

Condition monitoring is a booming research field, but condition monitoring of hydraulic components has not yet garnered the support of manipulator OEMs. These condition

monitoring systems have traditionally been seen as expensive features. However, since the level of automation in hydraulic manipulators on mobile machines is rapidly increasing, with more and more sensors are being added to these manipulators, and because companies are competing with intelligent manipulator features, condition monitoring is no longer merely an expensive feature. On the contrary, condition monitoring systems can be valuable assets to companies which are looking to stand out in the intensively competitive market sector. Moreover, as hydraulic manipulators are continuously developing into more complex machines, manual fault diagnosis is becoming increasingly more difficult. Condition monitoring solutions for hydraulic heavy-duty manipulators are also generally lacking in the literature.

These systems, that may be implemented locally or remotely, can offer valuable information about a machine's health to support fault diagnosis. One of the primary benefits of condition monitoring systems is preventative maintenance which entails early detection and diagnosis of slowly developing faults before they precipitate into undesirable system failures. These benefits can also be added to the list of generally acknowledged benefits of condition monitoring systems [7]:

- decreased downtime and lower operational costs, thus better machine availability;
- improved system performance and safety;
- reduced repair time through automatic fault diagnosis;
- improved system operational knowledge;
- minimised revenue loss and
- maintained product quality and customer relationships.

As a further motivation for hydraulic component condition monitoring, these systems could be highly important to the supervision of future autonomous machines and possibly a principal means of ensuring the proper functioning of manipulators with minimal human supervision. Concerning both autonomous and non-autonomous machines, the Industrial Internet (the Internet of Things, IoT) concept could offer a networking solution to enable hydraulic machine monitoring centrally on remote locations, instead of locally on the individual machines. This could also have major consequences for the effectiveness of monitoring for hydraulic manipulators because statistical analysis of normal and faulty systems could be carried out from the data sent over the network. Awareness about the possibilities of sensor network technologies and their benefits has increased overall, which should encourage future implementation of condition monitoring systems.

1.3 Motivation for Energy-Optimal Redundancy Resolution

Heavy-duty hydraulic manipulators in the forestry and construction industries are still open-loop controlled, but the industries' OEMs are interested in broadening their market offerings through introducing robotic closed-loop control to their machines to automate repetitive tasks. Löfgren [8] estimated that work productivity could be increased in the forest industry if repetitive tasks were automated. Furthermore, the mental and physical stress on the operator caused by simultaneous precision control of several actuators

via joysticks could be reduced. Robotic functionality might also consequently improve safety by enhancing the operator's alertness, however, obviously after the safety of the closed-loop control has been validated.

Robotic control of hydraulic manipulators is enabled by precision motion control of the end-effector, which requires resolving the individual joint trajectories of the manipulator to obtain the controller motion references. Regarding non-redundant hydraulic manipulators, these joint trajectories can be resolved through inverse kinematics. However, the typical hydraulic manipulators in the forest industry, for example, are equipped with a redundant telescoping link to increase reach and to ease material manipulability. For these redundant manipulators, kinematic inversion is non-trivial. There are, in fact, infinite joint motion trajectory solutions inside of the workspace that would satisfy the prescribed end-effector path.

The redundancy problem, whose resolution is obligatory for automation of forestry manipulators, for example, comes with the advantage of the possibility of joint trajectory optimisation. This possibility can be used to reduce energy consumption by resolving the joint trajectories to minimise hydraulic energy consumption over the end-effector path. Minimisation of hydraulic energy implies that the fuel consumption and diesel emissions of the machine can be reduced. Although minimum-time and conventional solutions, which may consume significant amounts of energy, have been proposed in the literature, an energy-optimal control option has not been considered at multiple valve-controlled hydraulic actuators and the hydraulic power system interaction level. This type of energy-optimal solution might also increase the wear-life of the manipulator and its hydraulic components, e.g., compared with minimum-time solutions. Finally, with ever-tightening emission standards for non-road mobile machinery arising from well-known environmental concerns, decreasing the energy consumption of hydraulic systems is highly important.

1.4 Research Methods and Restrictions

This research work begins by designing model-based condition monitoring for a hydraulic manipulator's valve-controlled cylinder actuator subsystem. Statistical and empirical residual thresholds are employed to incorporate a degree of robustness against modelling errors. This part of the research work deals with faults originating in hydraulic cylinders, valves and sensors. Common leakage and valve spool position jamming faults are extensively studied. Hydraulic fluid contamination and pump faults, however, are omitted. Software fault monitoring is also excluded (see the survey article [9]). The condition monitoring methods are experimentally validated in test-beds which are controlled by using a hydraulic proportional valve or a more complex mobile hydraulic valve. Statistical condition monitoring methods, such as logistic regression and correlation analysis, are beyond the scope of this work.

Model-based condition monitoring methods are generally chosen because non-model-based methods may suffer from poor fault sensitivity and fault cause identification capability. The condition monitoring solution based on a nonlinear Kalman filter was chosen in publications **P-I**–**P-III** to extend existing, well-received model-based condition monitoring solutions based on nonlinear Kalman filters to hydraulic systems with varying cylinder load. In publication **P-IV**, a non-Kalman filter-based solution was selected for the load-independent condition monitoring of mobile hydraulic systems because the

modelling errors of pressure-compensated, two-stage hydraulic proportional valves were somewhat more significant than the modelling errors of the simpler hydraulic proportional valves studied in publications **P-I**–**P-III**. The proposed relatively accurate statistical modelling of modelling errors was also easier to implement for the simpler model-based condition monitoring solution than for the nonlinear Kalman filter-based solution. These reasons led to applying the Kalman filter to the condition monitoring solutions presented in publications **P-I**–**P-III** and for choosing the simpler model-based condition monitoring solution presented in **P-IV** over other options.

The redundancy resolution part addresses the reach redundancy of common 3-DOF and 4-DOF hydraulic manipulators. Other manipulator types are not considered because they are not representative of industrial heavy-duty manipulators. The redundancy resolution proposed in publications **P-V** and **P-VI** is hydraulically energy-optimal along a prescribed path and covers the typical LS and CP systems and enforces actuator position, velocity and acceleration limits. Solving redundancy with respect to non-energy objectives and along non-prescribed paths is excluded from this work to limit the scope. Redundancy optimisation is performed chiefly using dynamic programming (DP) [10]. For comparative purposes, we also resolve the redundancy using standard direct optimisation tools [11] and pseudo-inverses [12, 13]. Indirect optimisation tools are not employed. Mathematical modelling is employed for defining manipulator kinematics, dynamics and the cost functions. The design of the required closed-loop control is only briefly discussed within the scope of this thesis. The proposed redundancy resolutions are validated through simulations and experimental work.

1.5 Thesis Contributions

As the main contributions of this thesis, novel condition monitoring solutions and energy-optimised redundancy resolution are proposed as operational improvements for hydraulic manipulators to facilitate robotic control of these manipulators. The publication contributions can be summarised as follows:

P-I: A systematic parameter sensitivity identification procedure was proposed for improving modelling accuracy and consequently the reliability of model-based condition monitoring. The most sensitive parameters in the proportional valve-cylinder system were analysed exploiting global sensitivity analysis (GSA) [14] for the first time [15].

P-II: A load-independent model-based condition monitoring algorithm based on the unscented Kalman filter (UKF) framework was designed for the proportional valve-controlled cylinder actuator subsystem on hydraulic manipulators. The fault detection capability of the algorithm was experimentally validated, and new fault isolation patterns concerning the spool position jamming of the hydraulic valve were identified [16].

P-III: Analysis of the experimental results confirmed the possibility of replacing high-cost joint angle encoders with low-cost (MEMS) motion sensors in the UKF-based condition monitoring algorithm [17].

P-IV: A model-based condition monitoring algorithm was designed for the mobile hydraulic valve-cylinder subsystem on hydraulic manipulators. Methodology

was developed to design adaptive residual thresholds based on statistical modelling errors. The fault detection capability of the algorithm was experimentally validated, and fault isolation patterns were identified [18].

P-V: A dynamic programming (DP)-based redundancy resolution algorithm was designed for minimising the hydraulic energy consumption of the LS and CP systems. System dimensionality was significantly reduced by exploiting redundancy to optimise only the extension cylinder motion. Joint limits were guaranteed (including the position, velocity and acceleration). Efficacy of the proposed procedure was demonstrated on a commercial manipulator model and through closed-loop simulations [19].

P-VI: The DP-based redundancy resolution algorithm was extended to encompass complicated 4-DOF forestry hydraulic manipulators. The method's energy-optimality was evaluated against the competing redundancy resolutions in a heavy-duty manipulator with CP and LS systems. Actuator position, velocity and acceleration bounds were satisfied in the solution [20].

1.6 The Author's Contribution to the Publications

P-I: The author wrote the paper and developed the approach for verifying condition monitoring models using GSA. Professor Jouni Mattila, the academic supervisor, reviewed the paper and suggested major improvements.

P-II: The author wrote the paper and developed the FDI framework. The author carried out experimental and simulation work which revealed the novel fault patterns. Professor Jouni Mattila, the academic supervisor, reviewed the paper and suggested major improvements.

P-III: The author wrote the paper, excluding Section II, and performed a wide range of experiments to facilitate the comparison between the micro-electromechanical systems (MEMS) sensor system and the encoder sensor system. Professor Jouni Mattila and the co-writers responsible for writing Section II, Dr. Juho Vihonen and Dr. Janne Honkakorpi, reviewed the paper and suggested major improvements.

P-IV: The author wrote the paper and developed the methodology for the system model and adaptive thresholds to facilitate fault detection for a mobile hydraulic valve system. Professor Jouni Mattila reviewed the paper and suggested improvements.

P-V: The author wrote the paper and designed the problem formulation and the energy-optimal redundancy resolution for a hydraulic manipulator. The author performed simulations to verify the redundancy resolution presented. Professor Jouni Mattila reviewed the paper and suggested improvements.

P-VI: The author wrote the paper and expanded the proposed redundancy resolution to complex commercial manipulators. The author carried out experiments to evaluate the performance of the redundancy resolution. Professor Jouni Mattila reviewed the paper and suggested improvements.

1.7 Outline

This introductory part of the thesis is divided into five sections. The contents of the remaining sections are summarised below.

Chapter 2 is a review of state-of-the-art in model-based condition monitoring of cylinder-valve actuator subsystems as found on hydraulic manipulators. The design principles of model-based condition monitoring systems are covered, and the typical hydraulic systems faults are briefly discussed.

Chapter 3 is a review of state-of-the-art in the redundancy resolution and optimal control of hydraulic manipulators. Furthermore, the numerical methods, that can generally be used to solve optimal control problems, are introduced and discussed.

Chapters 4 and 5 summarise the results and discoveries presented in the individual publications, comprising this research work.

Chapter 6 provides the conclusions and future work suggestions.

The research publications **P-I**–**P-VI** are appended at the end of the thesis.

2 State-of-the-Art in Condition Monitoring of Hydraulic Systems in Hydraulic Manipulators

In the introduction, the motivation for the implementation of condition monitoring systems on mobile hydraulic manipulators was described. These condition monitoring systems, for example, enable quicker detection of the faults in the manipulator's hydraulic systems. Condition monitoring solutions for hydraulic components, thus, have been developed in the literature. These solutions are increasingly model-based because the classical methods of trend or limit-checking of measurable outputs do not usually allow for sophisticated fault diagnosis [21]. The substantial rise in computational power from embedded processing units in recent years facilitates the implementation of model-based condition monitoring solutions on hydraulic manipulators. Embedded sensor technology available on mobile hydraulic systems has also recently grown to support this development (see e.g. [22], the website of a commercial OEM of magnetostrictive in-cylinder piston position sensors).

This section covers state-of-the-art solutions in condition monitoring of hydraulic manipulators, with a focus on model-based hydraulic cylinder-valve systems. We first classify condition monitoring terms in Section 2.1 and follow with an introduction to the principles, design tasks and methods of model-based fault detection in Section 2.2. The faults of hydraulic systems are discussed in Section 2.3. Finally, in Section 2.4, state-of-the-art condition monitoring solutions applicable to hydraulic manipulators are reviewed, with an intention to demonstrate the solutions' strengths and weaknesses.

2.1 Fault Detection and Isolation

Condition monitoring is founded on the fault detection and isolation (FDI) principle. FDI comprises two-steps, in which recognition that a fault has occurred (i.e. fault detection) is followed by fault cause localisation (i.e. fault isolation) [23–25]. Fault detection and diagnosis (FFD) and fault diagnosis (FD) are commonly used as synonyms for FDI [26, 27]. Fault size identification is also possible, but it is not discussed in this thesis [28].

Fault detection methods can be categorised in a number of ways: the basic conceptual categorisation is into the model and non-model-based methods (see Figure 2.1). These model-based methods require process knowledge, thus a system model; however, time-domain correlation analysis, frequency domain spectral analysis and wavelet analysis,

which are examples of non-model-based methods, do not require a system model. Some authors include spectral methods in signal model-based methods [29].

Fault isolation methods are sometimes categorised into inference and classification methods (see Figure 2.2). Inference methods require prior knowledge about symptom-causalities, whereas classification methods do not [21, 27]. Regarding model-based methods, fault isolation methods can also be grouped based on residual design principles into directional and structured residual approaches [30]. Generally, the terminology for fault isolation methods is slightly ambiguous, because different FDI textbooks use different terms [21, 28].

Non-model-based methods are often simpler than model-based methods, but non-model-based methods may suffer from poor fault sensitivity and fault cause identification capability. We focus exclusively on model-based methods, which should have better performance in these areas.

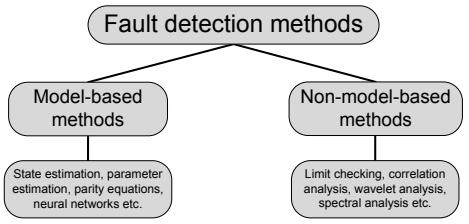


Figure 2.1: Fault detection methods

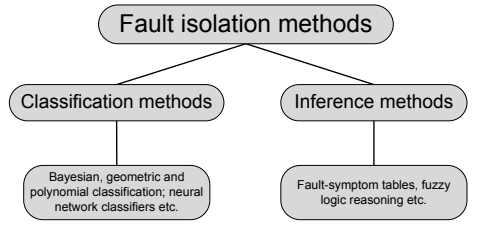


Figure 2.2: Fault isolation methods

2.2 State-of-the-Art in Model-Based Condition Monitoring

The principles of model-based condition monitoring are covered in this section, and model-based methods are briefly classified. The design tasks, which need addressing to improve robustness, are also discussed.

Principles

Fault detection in model-based methods is fundamentally enabled by the computation of the difference or residual between the system and the model output, as shown in Figure 2.3 (residual generation). This residual, ideally, deviates from zero only when the system is faulty.

The problem from the viewpoint of fault detection is that, in practice, this residual feature is non-zero also in nominal operating conditions because model residuals are susceptible to various sources of modelling errors and time-dependent disturbances. For this reason, FDI methods which are insensitive to modelling errors (unknown dynamics, unknown inputs and non-fault-related parameter variations) and sensitive to faults have been designed. By using these methods, robustness against non-zero residuals in nominal conditions is ensured through reliable residual generation or robust residual thresholds, which define the alarm thresholds. In this research work, robust residual thresholds were designed. Because of the importance of increasing solution robustness, the major design tasks of model-based condition monitoring are discussed next.

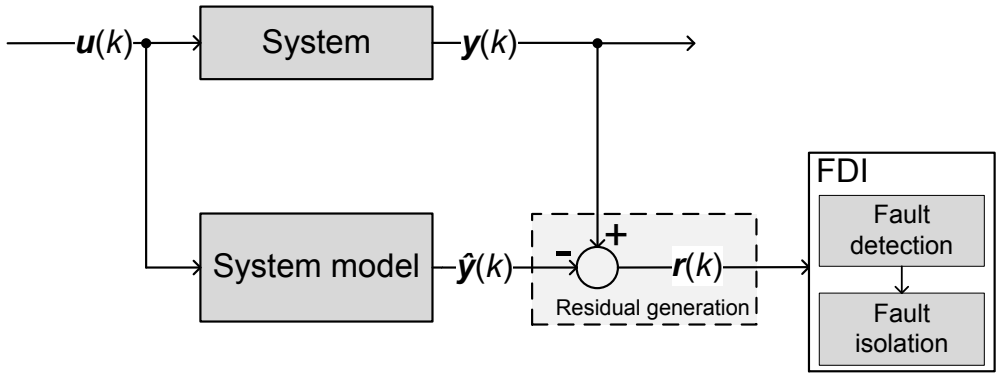


Figure 2.3: Model-based condition monitoring is founded on the principle of analytical redundancy. The fundamental problem is reliable residual generation or the design of robust residual thresholds, which enable reliable fault detection and isolation; adapted from [31].

Design of Model-Based Fault Detection

The first design task is to create a system model and identify its parameters. There are plenty of excellent sources in the literature for the modelling of physical hydraulic systems; see for example the classic source [32], as well as [33] and [34]. The most sensitive model parameters can also be determined at this stage, and they can be fine-tuned to improve modelling accuracy. This fine-tuning was carried out for the proportional hydraulic valve system based on GSA [15, 16].

After the model dynamics are verified against the measured dynamics, the second design task is to increase fault detection reliability by attenuating the effect of the modelling uncertainties and errors on the residual. This can be achieved by robust residual generation or residual thresholds. Residual evaluation methods are also applied to achieve the robustness goal.

For robust residual generation, the survey paper [9] suggests "designing a residual generator that is of high sensitivity for faults and has robustness to model uncertainty" and proposes design approaches to achieve this goal. The proposed first approach deals with the exploitation of knowledge about modelling uncertainty in the design of a robust residual generator. This approach may produce a desirable outcome if the residuals can be set arbitrarily close to zero in nominal conditions; however, this approach may be impractical because the modelling uncertainty dynamics are usually unknown. In those cases, when a suboptimal approximation of the modelling uncertainty by a suitable probability distribution transforms the residual into a zero-mean white noise sequence under nominal conditions, the approach is justified. The second proposed approach involves solving an optimisation problem to find the extreme value of the residual performance index to minimise the rate of false alarms and to maximise the rate of fault detection. These robust residual generation approaches were not pursued further.

The selection of appropriate residual evaluation functions is important for determining whether the system is faulty or not, and it may enhance fault detection reliability (i.e., decrease the occurrence of false alarms). The common residual evaluation functions include filters such as the moving average filter (an approach utilised in this research work); the sequential probability ratio test (SPRT), which computes the log-likelihood

ratio to investigate the fault likelihood of stochastic signals; and the generalized likelihood ratio test (GLRT), which is a convenient test if the residual is a Gaussian white noise sequence. One can also use the generated residual as it is, but usually these time-varying residuals are statistically processed to improve fault sensitivity and to reveal slowly developing faults. Furthermore, multiple residuals can be transformed into a single residual, for example, by using the (windowed) root mean square norm [30, 35, 36].

Most often, the residual evaluation functions described above cannot effectively eliminate the possibility of false alarms. Thus, fault detection reliability is enhanced by defining fixed (constant) or adaptive (dynamic) thresholds on the residuals. When the residual exceeds the threshold or, for the case of the separate positive and negative thresholds, is outside of these thresholds, an alarm is given to indicate the presence of a fault.

Whilst fixed thresholds remain stationary, adaptive thresholds are designed to vary depending on the system states and inputs. Fixed thresholds are sub-optimal because the residual sequence is usually non-stationary. Adaptive thresholds, in contrast, are generally more optimal than fixed thresholds because the modelling error dependence on the system states and inputs can be considered through statistical or empirical models. Although adaptive thresholds are practical, the fault-sensitivity achieved with the approach may decrease when the thresholds do not precisely track the residual. The statistical threshold design approach, however, seems to be a promising method for enhancing fault detection capability.

Model-Based Condition Monitoring Methods

There are multiple ways to categorise model-based condition monitoring methods, for example, by the way residuals are generated or faults are otherwise detected. Frank [9, 24] divides model-based condition monitoring methods into the following categories:

1. State observers,
2. Parameter estimation,
3. Parity equations and
4. Neural networks.

State Observers

The state observer approach in its various forms is highly popular and includes Kalman filters, unknown input observers, disturbance observers, nonlinear observers, linear observers, reduced-order observers (known as the output observer), full-state observers, H_∞ filters and diagnostic observers, amongst others [37]. The state observer approach works by reconstructing the state or at least the measurable outputs by utilising the available outputs and, in some cases, by exploiting knowledge about the modelling errors, measurement noise or other disturbances. Fault detection is enabled by model residual analysis.

The typical non-robust, non-stochastic (Luenberger), full-order linear observers are designed as follows [21]:

$$\begin{aligned} \dot{\hat{\mathbf{x}}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{L}\mathbf{r}(t) \\ \mathbf{r}(t) &= \mathbf{y}(t) - \mathbf{C}\hat{\mathbf{x}}(t) \end{aligned} \tag{2.1}$$

where $\mathbf{r}(t)$ is the residual vector. The gain $\mathbf{L}$ is chosen to stabilise the system, i.e. $(\mathbf{A} - \mathbf{LC})$ needs to be a stable matrix (Hurwitz), which conveniently forces the state error $\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)$ to zero in the absence of disturbances, modelling errors and faults. We can also design a reduced-order observer, which estimates the reduced state vector $\mathbf{z}(t)$ derived from $\mathbf{x}(t)$ by using the linear transformation $\mathbf{T}$ on $\mathbf{x}(t)$. Most observers are generalisations of this typical observer structure defined in Eq. (2.1). Nonlinear observers are also based on the same principles, although they generally have nonlinear, non-constant gains. This complicates the system analysis, because the linear matrix analysis techniques do not apply.

Kalman filters are the optimal stochastic state observers, and they minimise the mean squared state error (MSE) to optimally fuse model knowledge with noisy measurements. This minimisation usually requires that the state and measurement noises are additive. The state estimates obtained from a Kalman filter are generally free from measurement noise and state (process) noise that represents the model uncertainties. When applied to FDI, Kalman filters require that the process and measurement noise covariances are selected so that the estimation performance is balanced in the presence of modelling errors and faults. When the filter is designed correctly, the Kalman filter estimate does not diverge from the true state, and the residuals remain sensitive to faults. In practice, when the modelling errors are significant, the residual fault sensitivity suffers. Nonlinear versions of the Kalman filter include the extended Kalman filter (EKF) and the UKF, amongst others [38]. The UKF has been used in this research work when diagnosing the system comprising a proportional hydraulic valve and a hydraulic cylinder [16, 17].

Robust observers are designed to be independent of the unknown inputs. An example robust observer is the unknown input observer (UIO). This UIO requires knowledge about the structure of the unknown input to facilitate the selection of an unknown input distribution matrix; hence it may be difficult to design. These approaches, although theoretically promising, have not been further considered in this thesis.

Parameter Estimation

Fault detection using parameter estimation is based on estimating the system's health using fault-sensitive model parameters. Hence, faults can be detected by comparing these estimated parameters to nominal system parameters, and residuals need not be generated. In the approach, a fault may be described by using an additive term which acts on the system parameter [28]. The parameter estimation approach can yield detailed component health information because this approach ideally allows for multiple parameters to be simultaneously estimated from a limited number of system measurements. The major problem of the parameter estimation approach is the parameter estimation reliability when persistent excitation from the inputs and the system state cannot be guaranteed and when the nominal model parameters are unknown, for example when the parameters are nominally time-variant. Regarding hydraulic systems, the effective bulk modulus is an example of a difficult-to-estimate time-variant parameter, varying with the pressure level and piston position changes.

Parity Equations

The parity equations approach is a simplified version of the observer-based approach in the sense that the system outputs are not actively used to correct the model predictions to increase the estimation reliability. The fundamental idea of the parity approach is to establish a verified parity between the model output and the system measurement, that enables the computation of the residual signal, as shown in Figure 2.3. The parity equation approach is amongst the simplest methods, although a plant model transformation can be used.

Neural Networks

The neural network approach is an alternative modelling approach, that enables modelling by black-box parametrisation. As a result of its black-box nature, the model parameters of the neural network have no physical meaning. This means that when the actual parameters change, updating the neural model is non-trivial, whereas updating a physics-based model is trivial. For example, consider the hydraulic manipulator model whose several cylinder-valve pairs have similar dynamics but different valve and cylinder size parameters. The neural network model would have to be individually trained for each valve-actuator pair, whereas parameter changes to the physics-based model would often be sufficient. The neural network modelling approach, however, is useful when the physics-based models are too complex, inaccurate or computationally inefficient. Neural network approaches might become popular and relevant in future applications following the rapid development and use of deep learning techniques in other areas of engineering.

Other Classifications

Hwang [30] divides the model-based methods in FDI into the following higher number of categories: 1) full-state observer-based methods, 2) unknown input observers, 3) parity relations approach, 4) optimisation-based approach, 5) Kalman filter-based approach, 6) stochastic approach, 7) system identification approach, 8) nonlinear systems approach, 9) discrete event system/hybrid system approaches and 10) artificial intelligence techniques. Other classifications and broad discussions of the main methods behind model-based FDI can be found in textbooks [21, 28, 29, 36, 39, 40].

2.3 Model-Based Fault Detection of Hydraulic Systems

The problems concerning the fault detection of hydraulic systems are largely related to treating modelling errors effectively. Specific characteristics, that complicate the development of mathematical models for the FDI of hydraulic systems and may induce those modelling errors, include dynamic nonlinearities (e.g. flow, friction and pressure dynamics), general complexity and parameter variations in hydraulic systems.

The nonlinearities affecting hydraulic system dynamics imply that linear fault detection observers cannot be used without resorting to linearisation of nonlinear hydraulic equations. This linearisation may not be practical because it can yield high modelling errors due to wide operating pressure ranges and hydraulic valve discontinuities. Moreover, the hydraulic systems' complexity can complicate the modelling of hydraulic systems in general, which

tends to increase the likelihood of modelling errors. Hydraulic systems are well-known to be more difficult to model than electrical systems (see [41, 42]). Furthermore, the operation of hydraulic systems can be affected by parametric noise, and system responses are usually sensitive to changes in operating conditions. Time-varying parametric changes, which can alter system dynamics and complicate modelling endeavours, are quite common and often occur as a result of fluid temperature fluctuations, component wear and oil cleanliness changes. These characteristics of hydraulic systems should be considered in order to design reliable model-based FDI for hydraulic systems.

In this thesis, we treat the nonlinearities by using effective nonlinear hydraulic models. Furthermore, our model structures were formulated to increase modelling accuracy and to remove the dependency on joint and cylinder friction and the load of the hydraulic manipulator. This means that changes in friction parameters, for example, cannot be detected, which is an unfortunate side-effect of the proposed solutions. Adaptive thresholds for residuals were designed to attenuate the effect of modelling errors. Moreover, we used a moving average filter on the residuals. System parameter changes were not extensively considered in the scope of this research work. They were left for future work, which is discussed in the conclusions.

Hydraulic System Faults

Hydraulic systems can be affected by various faults, for example,

1. Hydraulic component faults (including leakages, erosion, jamming, ageing-related wear, friction, cracks, ruptures, fractures, loose parts and electrical faults) and

2. Sensor faults (offset, drift, hysteresis, dead-zone and electrical faults).

Research studies have shown that of the hydraulic system components the valves and actuators are the components most likely to break down (see [40] and [42]). Some of these valve and actuator faults include slowly-developing valve spool jamming (locking), ageing-related erosion of the valve spool's control edge, return spring failures, increased spool friction and sealing failures, that cause external or internal leakages in the hydraulic valve or actuator. Different types of sensor faults, typically in the position and pressure sensors, can also be paradoxically common and may affect hydraulic system availability.

Hydraulic component faults are often a result of contaminated hydraulic fluid when the filters fail to protect the system. The fluid can accumulate impurities, such as air, water, cloth fibers, dust particles and minute particles of worn-off metal or other material, during operation and assembly. These impurities alter the stiffness property of the system (bulk modulus), causing dynamical changes, which decrease the system's performance. Moreover, the debris in the fluid causes abrasion and erosion in the hydraulic valves due to the valve's minor clearances. There are indications that, in the short-term, the dirt particles can cause valve spool jamming and valve spool position offset, and leakages are probable in the long term; see [43, 44]. Contamination could also increase system noise levels. Based on the consequences and commonness of the hydraulic valve, cylinder and sensor faults, the focus of this research work on the detection and isolation of these component faults of heavy-duty hydraulic manipulators is well justified.

2.4 State-of-the-Art Fault Detection and Isolation Methods Applied to Hydraulic Systems

The field of FDI is, generally speaking, quite broad, and it has been developing since its inception in the early 1970s. However, this thesis considers a subset of model-based condition monitoring because we focus on condition monitoring of hydraulic systems on manipulators; hence, a fairly representative background review can be written. Specifically, reviewing the literature pertaining to model-based hydraulic system condition monitoring excluding the pump components is carried out in this section to keep in line with the objectives of this research work.

The state-of-the-art method in the field of model-based FDI, particularly in terms of fault sensitivity and practical applicability, includes the method introduced by Münchhof [45]. This method uses a combination of parameter estimation and parity equations, which are formed from physical principles and allow for the generation of residuals for the purpose of detecting and isolating a comprehensive set of sensor and valve faults. This method has a significant fault detection capability; for instance, sensor biases of 1% of the maximum deflection, chamber pressure offsets of 1 bar and piston displacement faults of less than 1 mm could be detected using this method. The method requires the supply pressure, cylinder chamber pressure, piston displacement, valve spool displacement and fluid temperature measurement. The downside of this method is that most of these measurements are not available in industrial systems; the spool displacement measurement can be particularly difficult to obtain.

The work also compared the physical model-based approach to a local linear model tree (LOLIMOT)-based neural network approach to determine the best candidate for FDI. The physical model provided the most accurate FD and was therefore chosen. A weakness of the proposed method is that removing the valve spool displacement measurement was not considered. Fault sensitivity and robustness with and without the spool measurement, hence, were not compared. It is likely that the method is more inaccurate without this measurement, although modelling of the valve spool dynamics could be feasible.

Unfortunately, a KF implementation was not compared with the method proposed. The test-bed system used was relatively simple (a fast-acting proportional valve drove a spring-loaded hydraulic cylinder). Robustness against modelling uncertainty was enhanced by using fixed residual thresholds and temperature-dependent valve coefficients. Hence, this method may not be applicable to complex hydraulic systems.

Münchhof [46] also described a multi-model approach for online FDI and a model adjustment approach for offline FDI. In the online approach, fault models and nominal models are run in parallel. The model with the lowest squared output error is seen to represent the mode of the system. Fault isolation of internal leakage, additive sensor bias and multiplicative sensor bias are possible. The same linear test-bed was used as in [45]. Experimental results show that the approach is not sensitive enough to distinguish additive sensor biases from multiplicative sensor biases.

In the offline approach, a fault-free model of the online approach is used, and its parameters are estimated one at a time using a one-dimension (1D) interval halving optimisation algorithm. In the experiments, a sensor fault size of 1 bar was estimated correctly with an accuracy of ± 0.5 bar. The advantage of the proposed methods is that no residual thresholds are required, but the disadvantage is that robustness against modelling errors

is not realised; thus, the methods rely on the existence of an accurate system model in all operating points.

The Kalman filtering approach has been popular in the field of fluid power to detect faults. For instance, An [47, 48] proposed a method using a fault-free EKF to detect actuator leakages under unknown external loading by pressure-state estimation. External leakages and internal leakages across cylinder chambers as minor as 0.25 L/min were detectable. The EKF was used to estimate the unknown external force. The approach was shown to be fault-sensitive with sinusoidal inputs and somewhat sensitive with pseudo-random inputs.

The authors have also designed a scheme based on Wald's method to quantify the size of the leakage [48]. In that paper, Wald's method was applied with the assumptions of Gaussian residuals and improbable incorrect fault detection. Contrary to the usual procedure of using a single hypothesis for determining a faulty condition, multiple hypotheses based on the mean averages of the residuals were formed to determine the leakage level. In these works, the systems considered were simpler than our mobile hydraulic systems; in the external force estimation, the presented approach could be insufficient.

The EKF has also been applied to FDI by Zavarehi [49], where a parameter estimation approach was taken to detect abrupt and incipient valve faults. The choice of the parameter estimation approach over a state estimation approach was motivated by the lack of quality measurements. This work is probably the closest to my research work, in particular because the experimental 1-degree-of-freedom hydraulic manipulator system was somewhat more complex than usual and the hydraulic valve used was a two-stage pilot-operated valve. A novel model for this system was constructed and validated. Fault detection was realised by applying the SPRT to the estimated parameters comprising the solenoid resistance, viscous damping, coulomb friction, flow force coefficient and orifice areas. Robustness was added by utilising neural network pattern classifiers in the fault decision process to diagnose the faults. The effective orifice areas, including the dead zones, were estimated online without prior knowledge of the valve orifice size or profile, and they were shown to be excellent fault predictors because of an accurate spool position measurement. The modelling and identification approach presented, overall, might be too complex for industrial application.

Chinniah [50] estimated the effective bulk modulus and viscous damping coefficient parameters using the EKF when applied to a hydrostatic, pump-controlled electro-hydraulic actuator. The choice to estimate these parameters was motivated by their fault-sensitivity. Most often, a decreased bulk modulus is the result of entrapped air in the hydraulic fluid, and the viscous damping coefficient can change as a result of seal wear or hydraulic fluid contamination. An iterative approach to parameter estimation was taken, meaning that the bulk modulus and the viscous coefficient were separately estimated. The estimated bulk modulus was an input to the viscous coefficient estimation. The study showed that accurate estimation of the bulk modulus requires a sufficient input frequency and that this frequency can be approximately found by using the sensitivity function derived from the system transfer function. However, a structured input with this specific frequency is not typically applied when the manipulator is typically commanded. Therefore, this method is usable offline only.

In a recent paper by Sepasi and Sassani [51], the UKF was applied to detect leakages and load changes in hydraulic system with a constant, known external force by means of cylinder chamber pressure estimation. The authors were able to detect and isolate leakage faults and load changes, but controlling the system with pseudo-random inputs

was not considered. The method is not applicable to hydraulic manipulators with varying loads. We used the UKF in this research work in the publications **P-II** and **P-III**. In our approach, the UKF was designed to be load-independent.

Chen [52] designed a linear Luenberger observer scheme to detect and isolate internal leakage and pressure sensor offsets in a pump-controlled system, which drove a hydraulic cylinder against a load cylinder. The possibility of decoupling the external force from the state equations was shown by considering the piston velocity as an input. This approach appears to be similar to what we used in this work. The linear observer used might not be suitable for valve-controlled systems because of the flow mode nonlinearity. In the considered system, because of modelling errors, the residuals did not converge to zero in the fault-free cases. The presented method was able to isolate and detect a 5-L/min-internal leakage and 20% sensor biases. This method was not particularly fault-sensitive, even though the experimental system was not highly complex.

Khan [53] applied a nonlinear observer to the problem of fault detection and combined that with Wald's sequential test. The performance of the observer was studied by changing the bulk modulus and viscous friction through simulation. These simulations showed that the observer could not converge to the measurements in the presence of these varying parameters. Therefore, this observer design, in practice, could not satisfy the requirements of reliable fault detection. Experimental tests were performed, which showed that a 26% decrease in supply pressure and a 30% bias were detectable.

Tan and Sepehri [54] estimated the parameters of a nonlinear Volterra model and geometrically detected and isolated internal and external leakage, incorrect supply pressure and fluid contamination. Experimental results on a linear test-bed were shown regarding the detection of incorrect supply pressure. Simulations were performed concerning the other faults. Judged based on the experiments, the method was not fault-sensitive; in particular, the false alarm rate was 15%, and faults were correctly isolated 80% of the time. The geometrical FDI method designed, however, was the key contribution of this article.

Le [55] proposed a neural network-based fault detection method, but this method suffered from weak fault sensitivity: only fluid leakages over 1 L/min were detectable. The fault sensitivity of this approach seemed inferior compared to the LOLIMOT network proposed in [45].

The UIO has been recently applied to hydraulic systems by Mondal [56], in which the UIO was designed based on linear matrix inequalities to be robust against noise and uncertainties and to detect parameter changes. The fault sensitivity of the observer was tested by introducing a 45% spring coefficient reduction to the spring resisting the movement of the hydraulic actuator. This fault was shown to be detectable and isolatable. The observer designed seems interesting, but unfortunately, only simulations were provided.

Garimella and Yao [57] proposed an adaptive robust observer, which is robust against modelling errors, to detect electro-hydraulic system faults. This nonlinear observer [58] combines state estimation with parameter estimation. The well-known least-squares algorithm was employed to estimate the system parameters online. Because of the algorithm chosen, the authors could set upper bounds based on known inequalities as the parameter thresholds. The scheme was able to detect a 2% decrease in the supply pressure. The simulations showed that the method can accurately estimate the parameters related to internal leakage, external leakage and bulk modulus faults. However, the method's

experimental performance and robustness were not considered. In another paper by the authors [59], the applicability of the scheme to sensor fault detection was demonstrated.

Yu [60] proposed a parameter estimation method to isolate faults which have a similar state-space direction. In the article, a bilinear observer, that had been proposed by the author in [61], was used. A specific set of faults were isolated using the observer or the parameter estimation. The scheme was validated through simulations by using a model of a hydraulic motor controlled by an electro-hydraulic servo valve. In the fault simulations, a 10% change was introduced to the hydraulic motor's and the pump's efficiency, pump inertia and viscous friction coefficient. These simulations showed that the faults could be isolated. Experimental results were not provided to validate the simulation claims.

In Wang and Syrmos [62], an interacting multiple model (IMM) approach based on a bank of EKFs was proposed. The EKFs were employed to estimate the states of a linear test-bed, which was controlled by a two-stage servo valve. The EKFs were also employed to estimate the torque motor's resistance and effective bulk modulus. Each of the EKFs corresponded to a different mode, representing a nominal or faulty mode. The FDI was enabled by the probability scheme employed in the IMM. The scheme's performance was analysed using simulations, in which a 50% increase in the torque motor's resistance was introduced and the bulk modulus was reduced by 90%. Experimental results would have been required to investigate the robustness of the FDI scheme designed.

In Gadsden [63], the authors used the IMM approach with a bank of smooth variable structure filters (SVSFs) to estimate the probabilities of the normal mode, leakage mode and friction mode. The SVSF is a stable filter, meaning that the state error remains bounded. The authors extended the SVSF filter to include a state error covariance estimation. The SVSF-IMM approach proposed was experimentally verified in an electro-hydrostatic system and compared against the SVSF-KF. In these experiments, the SVSF-IMM approach's performance was superior to the KF-IMM approach in the condition mode recognition task. In the presence of leakage, the leakage mode probability was estimated to be of class 60% to 80%, and it was less than 5% in the absence of leakage. The advantage of this IMM approach is that fault isolation follows straightforwardly from the faults modelled. However, sometimes the faults cannot be easily modelled.

Shi [64] developed an adaptive threshold for the linearised model of an electro-hydraulic system, comprising a flapper-nozzle-type servo valve and a closed-loop controlled asymmetric hydraulic cylinder. This nonlinear hydraulic system was linearised to reduce computational complexity. The application of the proposed adaptive threshold, however, was motivated by linearisation errors. The threshold was developed statistically based on a series of position step inputs. The residual analysis demonstrated that the residual was stationary Gaussian noise with an approximately constant mean and standard deviation at the steady-states, but at transients, the noise was non-stationary because of input-dependent modelling errors. Hence, the adaptive threshold was set to a constant value at the steady-states; otherwise, it was set to a value dependent on the step input and fluid temperature. The practicality of the adaptive threshold proposed was evaluated by experimenting with several fault cases in the test-bed. After a comparison to a fixed threshold, the authors concluded that the adaptive threshold scheme causes fewer false positives (false alarms) and false negatives (missed alarms).

It should be stressed that this adaptive threshold scheme as such may not work with the complex hydraulic manipulators considered. The method proposed requires a closed-loop system and preferably a non-varying external force. Moreover, the linearisation errors could be much larger in a complex hydraulic manipulator because of the wider operating

space. However, this paper has motivated us to design an adaptive threshold scheme for the hydraulic manipulator.

In the hydraulic manipulator, the modelling errors are realistically non-stationary processes, meaning that they cannot be described by using a fixed distribution. Modelling these modelling errors, particularly in these complex systems, would be difficult without resorting to statistical approaches or neural networks [65]. In this thesis, we have hence designed the adaptive thresholds experimentally and statistically as a function of the system states and inputs.

Summary

To summarise our main findings, only a handful of papers, most prominently [49], deal with condition monitoring of the hydraulic components used in heavy-duty manipulators. The method introduced in [49] is thorough, but at the same time it may be difficult to implement from an industrial standpoint. The work by Münchhof [45] is also amongst the state-of-the-art methods, but the method developed, possibly, works only with non-mobile hydraulic valves, since the method does not consider the complex valve functions, such as the pilot operation and pressure compensation. Furthermore, based on our literature review, many of the methods which have been designed for hydraulic systems have not been experimentally validated. This is a major weakness because hydraulic system models are often subject to modelling errors (e.g., due to system complexity). Accurate model-based condition monitoring methods, which facilitate industrial hydraulic manipulator application, should, thus, be developed. Motivated by this situation, we design model-based FDI approaches for heavy-duty manipulators, with some of the shortcomings of the methods presented in the literature ameliorated.

3 State-of-the-Art Redundancy Resolution of Hydraulic Manipulators

A manipulator is kinematically redundant when it possesses more joint degrees of freedom than the amount required to complete a task. This kinematic redundancy property is very common in general robots and in 3-DOF and 4-DOF industrial hydraulic manipulators that are prismatically redundant. A wide variety of redundancy resolutions thus have been developed in the literature [12, 66 – 70]. However, many of these resolutions according to the author's knowledge do not deal with objectives concerning hydraulic manipulators, including energy consumption and minimisation of the pump-flow rate. In contrast to existing work, we deal with these objectives in this thesis. Energy-optimised redundancy resolution is highly important for the automation of hydraulic manipulators.

In Section 3.1, a general optimal control problem is formulated for a 4-DOF hydraulic manipulator to introduce the background of the redundancy resolution. In Section 3.2, state-of-the-art redundancy resolutions applicable to hydraulic manipulators are discussed. At the end, in Section 3.3, we introduce and compare numerical methods that can be employed to solve the optimal control problem.

3.1 Redundancy Resolution of Hydraulic Manipulators

Kinematic redundancy makes manipulator control difficult because there are infinitely many joint trajectories inside of the workspace that satisfy a prescribed path. However, at the same time, a major upside arises from the control complexity in that the joint trajectories can be chosen to optimise an objective. Figure 3.1 demonstrates the general idea of kinematic redundancy resolution through a comparison of two kinematic redundancy resolutions or joint trajectory solutions that satisfy a prescribed path from point A to point B in a straight horizontal line. The joint kinematic configuration of the illustrated 4-DOF hydraulic manipulator is the same for both of the compared kinematic redundancy resolutions at point A, but as the end-effector of the manipulator moves in a straight line to point B the joints of the hydraulic manipulator are observed to track two vastly different kinematic trajectories. The illustrated kinematic redundancy resolution one is a joint trajectory solution that is energy-optimal at the valve-controlled hydraulic actuator and hydraulic power system interaction level, and the exemplified kinematic redundancy resolution two is a conventional pseudo-inverse solution.

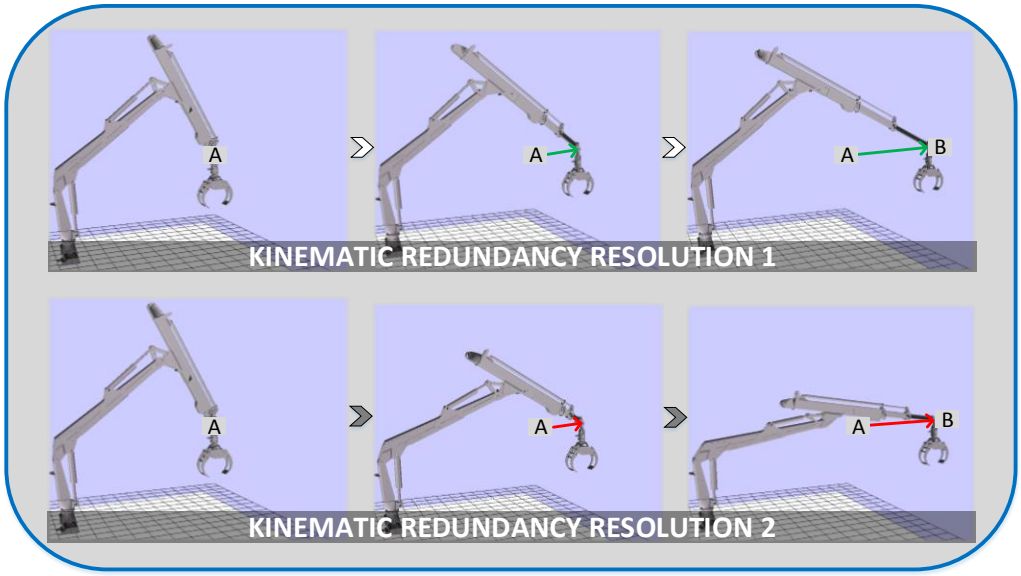


Figure 3.1: Two kinematic redundancy resolutions of a 4-DOF hydraulic manipulator satisfying a prescribed straight-line path from point A to point B.

Solving redundancy resolution, usually, begins with formulation of the optimal control problem, which encapsulates the optimisation objective and constraints. The aim is to solve the optimal control problem as effectively as possible. Redundancy resolution consists of the optimal control and state trajectories.

Redundancy Resolution Objectives

Kinematic redundancy can be exploited in many different ways, and it is often utilised to improve manipulator dexterity. This can help the manipulator manoeuvre around obstacles to reach specific workspace targets. In general, the following could be the objectives relevant from the perspective of hydraulic manipulators:

- minimisation of energy consumption,
- minimisation of task execution time,
- minimisation of motion jerk,
- obstacle avoidance,
- joint limit avoidance,
- joint force/torque minimisation,
- maximisation of load-carrying capacity,
- minimisation of the pump-flow rate,
- joint velocity minimisation and

- joint acceleration minimisation.

General energy-optimal redundancy resolutions have been proposed in previous work [69, 71]. Although these solutions are applicable to hydraulic manipulators, they may not offer the best performance from the perspective of the valve-controlled hydraulic system. This is a consequence of hydraulic power losses, which indicate that the energy consumption of the hydraulic system is generally higher than the energy consumption of actuators. These power losses, particularly, occur in multi-actuator hydraulic systems controlled by the typical valve-controlled LS and CP systems. Minimisation of hydraulic energy consumption, therefore, requires a redundancy resolution tailored for the hydraulic system.

Fundamental Problem and the Conventional Solution

At the velocity level, the fundamental problem we are trying to solve is the inversion of

$$\dot{\mathbf{x}}_\mathrm{t} = \mathbf{J}(\mathbf{q})\dot{\mathbf{q}} \tag{3.1}$$

to obtain the joint velocity vector $\dot{\mathbf{q}}$. Here, $\dot{\mathbf{x}}_\mathrm{t}$ denotes the desired end-effector velocity and the Jacobian matrix $\mathbf{J}(\mathbf{q})$ is the partial derivative of $\mathbf{x}_\mathrm{t}$ with respect to $\mathbf{q}$. At the acceleration level, the problem can be expressed equally as the inversion problem of

$$\ddot{\mathbf{x}}_\mathrm{t} = \dot{\mathbf{J}}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{J}(\mathbf{q})\ddot{\mathbf{q}}, \tag{3.2}$$

from which we would like to solve the joint acceleration vector $\ddot{\mathbf{q}}$. Here, $\ddot{\mathbf{x}}_\mathrm{t}$ denotes the desired end-effector acceleration and the time derivative of the Jacobian matrix is denoted with $\dot{\mathbf{J}}(\mathbf{q}, \dot{\mathbf{q}})$.

The main problem regarding kinematically-redundant hydraulic manipulators is that the Jacobian matrix $\mathbf{J}(\mathbf{q}) \in \mathbb{R}^{m \times n}$ cannot be easily inverted as it is non-square and it has fewer rows than it has columns. In other words, $n > m$, and there are more unknowns than there are equations. Hence, the inversion of the Jacobian matrix has an infinite number of solutions. A typical, kinematically-redundant hydraulic manipulator is presented in Figure 3.2.

A particular velocity-level inversion, and hence redundancy resolution, can be found by minimising the instantaneous weighted Euclidean norm of the joint velocities $\dot{\mathbf{q}}^\mathrm{T}\mathbf{W}\dot{\mathbf{q}}$ at the end-effector path as in

$$\min H = \min \frac{1}{2}\dot{\mathbf{q}}^\mathrm{T}\mathbf{W}\dot{\mathbf{q}} + \boldsymbol{\lambda}^\mathrm{T}(\dot{\mathbf{x}}_\mathrm{t} - \mathbf{J}\dot{\mathbf{q}}), \tag{3.3}$$

where $\mathbf{W}$ is an invertible diagonal weighting matrix, and $\boldsymbol{\lambda}^\mathrm{T}$ denotes the Lagrange multipliers required to satisfy Eq. (3.1). The joint velocities which minimise the cost H can be derived using multivariate optimisation theory [72]. Firstly, we have

$$\begin{aligned}
&\frac{\partial H}{\partial \dot{\mathbf{q}}} = \frac{1}{2}\dot{\mathbf{q}}^{\mathrm{T}}(\mathbf{W}+\mathbf{W}^{\mathrm{T}}) - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{J} = 0 \\
&\iff \dot{\mathbf{q}}^{\mathrm{T}}\mathbf{W} = \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{J} \\
&\iff \mathbf{W}\dot{\mathbf{q}} = \mathbf{J}^{\mathrm{T}}\boldsymbol{\lambda} \\
&\iff \dot{\mathbf{q}} = \mathbf{W}^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\lambda},
\end{aligned} \tag{3.4}$$

where matrix algebra [73] and the weighting matrix's property $\mathbf{W} = \mathbf{W}^{\mathrm{T}}$ have been used. Secondly, we have

$$\begin{aligned}
&\frac{\partial H}{\partial \boldsymbol{\lambda}^{\mathrm{T}}} = \dot{\mathbf{x}}_{\mathrm{t}} - \mathbf{J}\dot{\mathbf{q}} = 0 \\
&\iff \dot{\mathbf{x}}_{\mathrm{t}} = \mathbf{J}\dot{\mathbf{q}} \\
&\iff \dot{\mathbf{x}}_{\mathrm{t}} = \mathbf{J}\mathbf{W}^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\lambda} \\
&\iff \boldsymbol{\lambda} = (\mathbf{J}\mathbf{W}^{-1}\mathbf{J}^{\mathrm{T}})^{-1}\dot{\mathbf{x}}_{\mathrm{t}},
\end{aligned} \tag{3.5}$$

where the result of Eq. (3.4) has been used. By substituting the result of Eq. (3.5) into Eq. (3.4), finally, we get

$$\dot{\mathbf{q}} = \mathbf{W}^{-1}\mathbf{J}(\mathbf{q})^{\mathrm{T}}\big(\mathbf{J}(\mathbf{q})\mathbf{W}^{-1}\mathbf{J}(\mathbf{q})^{\mathrm{T}}\big)^{-1}\dot{\mathbf{x}}_{\mathrm{t}} = \mathbf{J}(\mathbf{q})^{\dagger}\dot{\mathbf{x}}_{\mathrm{t}} \tag{3.6}$$

as the joint velocities which minimise H. Here, $\mathbf{J}(\mathbf{q})^{\dagger} = \mathbf{W}^{-1}\mathbf{J}(\mathbf{q})^{\mathrm{T}}\big(\mathbf{J}(\mathbf{q})\mathbf{W}^{-1}\mathbf{J}(\mathbf{q})^{\mathrm{T}}\big)^{-1}$ is called the weighted pseudo-inverse of the Jacobian matrix, which, for example, leads to the instantaneous minimum Euclidean norm of the joint velocities as in $\dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{q}}$ when the weighting matrix $\mathbf{W}$ equals the identity matrix $\mathbf{I}$. Variations of this weighting approach have been presented (e.g. weighting with the manipulator inertia matrix minimises the instantaneous kinetic energy). The pseudo-inverses have also been derived in actuator coordinates [13].

These pseudo-inverse redundancy resolutions have severe practical limitations. Firstly, they lead to merely local, instantaneous minimisation of the objective function. Secondly, they can minimise only relatively simple objective functions. For these reasons, we propose an entirely different optimal control-based redundancy resolution to minimise the complex objective functions of the hydraulic manipulator and to produce globally optimal joint motions over the entire desired end-effector path. Here, we are interested in objective functions related to the energy consumption of typical valve-controlled hydraulic systems instead of objective functions which, for example, merely minimise the velocities of the joints. We also aim to satisfy the motion constraints of the hydraulic cylinders. There is much untapped optimisation potential on the kinetic level in redundant hydraulic manipulators, as can be seen from the two vastly different workspaces reachable with and without the redundant joint in Figure 3.3.

Optimal Control Problem Formulation

Effective energy-optimal redundancy resolution of a hydraulic manipulator requires resolving an optimal control problem tailored to the hydraulic system and subject to

Figure 3.2: 4-DOF hydraulic manipulator on a forest forwarder [74].

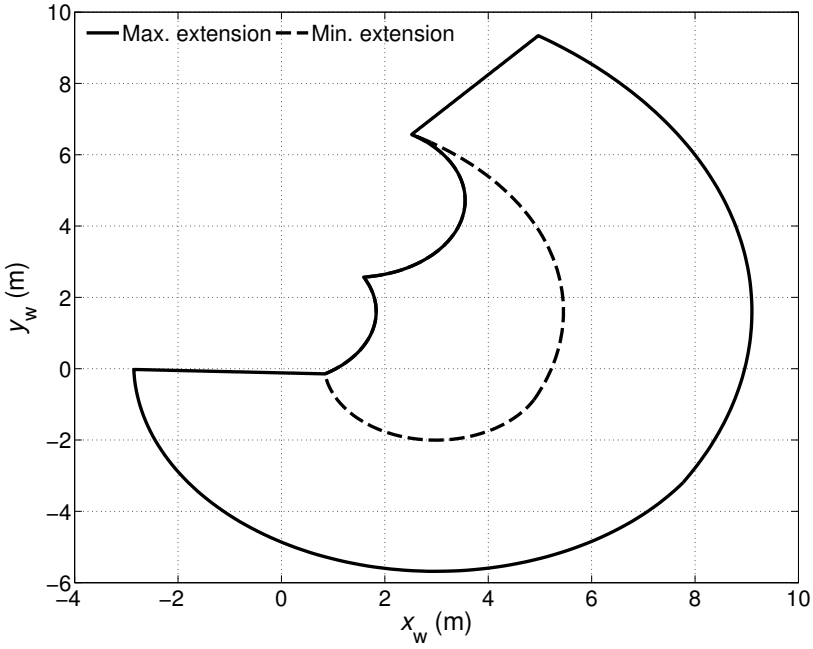


Figure 3.3: Typical theoretical reachable workspace of a 4-DOF hydraulic manipulator in the vertical plane. Maximum reachable workspace with fully extended extension joint and minimum reachable workspace with fully retracted extension joint.

many motion constraints. Since the hydraulic manipulators are complex with multiple actuators, it is pragmatic to simplify this optimal control problem. This generally means omitting hydraulic system dynamics. In continuous-time, the energy-optimal control problem can be formulated as follows:

$$\min_{\mathbf{u}\in\mathbb{U}} \int_0^{t_\mathrm{f}} P\big(\mathbf{x}(t),\mathbf{u}(t)\big)\,\mathrm{d}t \tag{3.7}$$

subject to:

$$\begin{aligned}
&\mathbf{x}(0) = \mathbf{x}_0,\\
&\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t),\mathbf{u}(t)\big),\\
&\mathbf{g}_\mathrm{i}\big(\mathbf{x}(t),\mathbf{u}(t)\big) \leq 0,\\
&\mathbf{g}_\mathrm{e}\big(\mathbf{x}(t),\mathbf{u}(t)\big) = 0
\end{aligned} \tag{3.8}$$

where $P \in \mathbb{R}$ is the hydraulic power cost of the hydraulic system to be minimised; t denotes the time; t_f is a fixed terminal time; $\mathbf{u} \in \mathbb{R}$ denotes the control vector; $\mathbb{U}$ is the control policy search space of feasible $\mathbf{u}(t)$, defined for $t \in [0, t_\mathrm{f}]$; $\mathbf{x}_0$ stands for the initial system-state vector; $\mathbf{x}$ denotes the system state vector; $\mathbf{f}$ denotes the system dynamics; $\mathbf{g}_\mathrm{i}$ includes the inequality constraints and $\mathbf{g}_e$ contains the equality constraints. These time-varying equality constraints $\mathbf{g}_\mathrm{e}$ can be used to enforce end-effector path tracking. Also, the state and control vector can be limited by using the time-varying inequality constraints $\mathbf{g}_\mathrm{i}$. These limits can include the position, velocity and acceleration limits. For hydraulic solution optimality, these limits must be defined in the actuator space. The velocity limits dominate because the maximum pump-flow rate is restricted, based on the pump dimensions. Furthermore, the piston positions are constrained by the physical limits of the cylinders. Piston acceleration limits, on the other hand, are not usually amongst the dominant limits because the hydraulic cylinders are capable of rapidly producing large forces. Hence, constant acceleration limits can be used.

If the state vector $\mathbf{x}$ (inefficiently) contained the 4-DOF manipulator's lower-order joint motion states such that

$$\mathbf{x} = [q_1\ \dot{q}_1\ q_2\ \dot{q}_2\ q_3\ \dot{q}_3\ q_4\ \dot{q}_4]^\mathrm{T} \tag{3.9}$$

then the control vector $\mathbf{u}$ should include higher-order joint states, i.e.

$$\mathbf{u} = [\ddot{q}_1\ \ddot{q}_2\ \ddot{q}_3\ \ddot{q}_4]^\mathrm{T}\,, \tag{3.10}$$

where q_i, for $i \in \{1, 2, 3, 4\}$, are the manipulator's generalised joint coordinates and the dot notation denotes time differentiation. However, it is computationally much more efficient to exploit kinematic redundancy and inverse kinematics to significantly reduce system dimensionality. This approach leads to reducing the state vector to

$$\mathbf{x} = [q_4\ \dot{q}_4]^\mathrm{T} \tag{3.11}$$

and the control vector to

$$\mathbf{u} = \ddot{q}_4. \tag{3.12}$$

Hence, we may reduce the problem size from eight to two states and four to one control. Since we should still obtain the remaining states and controls for the objective function, we can exploit the non-singular inverse kinematics which would exist without the redundant joint. At the position level, we can compute the remaining joint positions $\mathbf{q}^{(\mathrm{r})} = [q_1\ q_2\ q_3]^{\mathrm{T}}$ from

$$\mathbf{q}^{(\mathrm{r})} = [h_1\ h_2\ h_3]^{\mathrm{T}} \tag{3.13}$$

where h_1, h_2 and h_3 denote the inverse kinematic functions. At the velocity level, we can obtain the remaining joint velocities $\dot{\mathbf{q}}_k^{(\mathrm{r})}$ from

$$\dot{\mathbf{q}}^{(\mathrm{r})} = \left[\mathbf{J}_1(\mathbf{q})\ \mathbf{J}_2(\mathbf{q})\ \mathbf{J}_3(\mathbf{q})\right]^{-1} \left(\dot{\mathbf{x}}_{\mathrm{t}} - \mathbf{J}_4(\mathbf{q})\dot{q}_4\right) \tag{3.14}$$

because the Jacobian inverse matrix $\left[\mathbf{J}_1(\mathbf{q}_k)\ \mathbf{J}_2(\mathbf{q}_k)\ \mathbf{J}_3(\mathbf{q}_k)\right]^{-1}$ is non-singular. Here, $\mathbf{J}_i(\mathbf{q})$ denotes the column vectors of the Jacobian matrix. At the acceleration level, we can proceed similarly. This approach along with the cost functions proposed are presented in much greater detail in our publications on the energy-optimal redundancy resolution of hydraulic manipulators [19, 20].

3.2 State-of-the-Art Redundancy Resolution Applicable to Hydraulic Manipulators

This section covers state-of-the-art redundancy resolutions and a selection of other optimal solutions. We first discuss the point-wise optimal solutions, then proceed to the optimal control solutions and finally provide a brief summary of the findings.

Point-wise Optimal Solutions

On-line point-wise optimal solutions are very popular redundancy resolutions [75]. These resolutions usually involve some form of inversion of the Jacobian matrix, which equals the infinitesimal change in the end-effector position when the joints are infinitesimally actuated. Whitney [66] proposed the first Jacobian-based redundancy resolution based on the pseudo-inverse of the Jacobian matrix. This pseudo-inverse solution is approximate in the sense that it minimises the sum of the squared joint velocities. Weighted pseudo-inverse variants are popular when it comes to avoiding joint limits. Chan [68] proposed a redundancy resolution where the elements of a diagonal weighting matrix were designed to grow towards infinity very close to the joint limits, thus forcing the joint motion to stop at the limits. Another weighted pseudo-inverse was proposed by Beiner and Mattila [13] in the actuator space. This solution weighted the Jacobian by the manipulator's inertia matrix to minimise the kinetic energy instantaneously. This actuator space solution is useful for hydraulic manipulators that have cylinder position feedback.

In the case of the basic pseudo-inverses, the downsides are the violations of the joint velocity and acceleration limits and the solution locality. This solution locality implies that the solution is sub-optimal over the entire workspace path. Thus, when considering our prescribed end-effector path, pseudo-inverse redundancy resolution is not effective. Moreover, Suh [76] compared the local and global optimisations and demonstrated that the optimal control solution is generally better in terms of cost-function value and even stability.

Another important solution amongst the point-wise optimal solutions is the generalised pseudo-inverse (i.e. the gradient projection method). It expands the pseudo-inverse by introducing an additive term, which projects into the null-space of the Jacobian. This null-space projection matrix term produces joint motion that leaves the end-effector position unaffected (self-motion). The null-space expansion term allows for the minimisation of joint-dependent cost functions. Liégeois [67], for example, optimised the null-space joint velocity to limit joint range. Hollerbach [77] also proposed a torque-limited acceleration-level solution.

Flacco [12] recently presented new pseudo-inverse algorithms which minimise the velocity and acceleration norms with respect to joint limits that are expressed on the velocity or acceleration level. This algorithm is based on saturating the joint velocities or accelerations at the null-space. The velocity-level solution satisfies strict joint position and velocity limits, but the acceleration level can also enforce the joint acceleration limits. These pseudo-inverse solutions share many of the disadvantages of the other pseudo-inverses concerning the automation goal of the hydraulic manipulator, for example, the locality of the optimisation.

Another local approach is Vukobratovic's redundancy resolution [71]. That paper presents a generalised pseudo-inverse solution which is designed for electrical and hydraulic manipulators from the perspective of actuator energy consumption minimisation. Joint velocity limits and joint ranges were enforced. The manipulator considered in the numerical examples was not the typical 3-DOF or 4-DOF hydraulic manipulator. Significant hydraulic power losses were also neglected, contrary to the work presented in this thesis.

Beiner [78] presented a solution, which minimised the cylinder force norm of the typical 3-DOF hydraulic manipulator whilst considering the cylinder strokes. An interesting analytical optimum was developed and was verified using standard optimisation software. Functional analysis showed that the lift cylinder arm must be maximised, the tilt cylinder should be fully extended, the extension cylinder must be fully retracted and the lift boom should be in a horizontal position to minimise the cylinder forces. The solution is not straightforwardly extendable to prescribed paths. Velocity limits were also omitted from the paper, which is a disadvantage because the pump-flow rate cannot exceed the limits determined from the component dimensions.

As for general solutions for redundant manipulators, Zhang [79] presented a quadratic programming framework. The author developed velocity and acceleration-level solutions which include a minimum torque norm, inertia-inverse weighted torque, minimum kinetic energy and a minimum velocity norm scheme. These methods are interesting because they can impose all the joint limits. A primal-dual dynamical solver based on linear variational inequalities was the solver suggested because the other solvers are usually computationally expensive. Simulations were carried out to validate the solver.

Zhang [80] also extended his framework to handle online obstacle avoidance and introduced cost functions, such as a repetitive motion planning index for performing cyclic tasks

and a minimum acceleration norm. All of the cost functions were expressed in quadratic form. In the paper, different online quadratic program solvers were briefly compared. Furthermore, Zhang [81] combined the minimum torque and velocity norm schemes into a weighted approach, which seemed to prevent the occurrence of high joint velocity and torque whilst imposing close to zero final joint velocities.

Because Zhang's solutions [79 – 81] satisfy all the relevant joint limits, they could be applied to hydraulic manipulators. The actuator space representation proposed in [13] could be combined with Zhang's approach. However, these solutions are local and are not formulated to minimise hydraulic energy consumption.

Optimal Control Solutions

Optimal control methods are widely used to derive optimal control policies that are optimal with respect to the objective and constraints over the entire time-horizon. These optimal control methods usually fare better than instantaneous methods. A disadvantage of optimal control methods compared with the pseudo-inverses is their higher computational cost, making them mostly suited for offline application.

However, optimal control policies can be stored in a database and retrieved from there when needed. The problem is the selection of the policies that should be stored in the database. This problem was addressed by Ortiz Morales [82] and Mettin [83] in their articles on log-loader trajectory planning and control. The authors concluded that a selective strategy for solution storing might work. This was evident because the path patterns recorded whilst professional human operators were operating the machine demonstrated a limited number of those initial workspace points from which the end-effector movements begin. Therefore, the joint trajectories could be pre-computed.

Minimum-Energy Solutions

Energy-optimal redundancy resolution of hydraulic manipulators is a major objective of this thesis. Amongst the well-known energy-optimal solutions is the method presented by Hirakawa [69], which describes a trajectory planning method for minimising the electrical energy consumption of joint-driving electric motors. This method was variational, that was augmented by using third-order B-spline curves to interpolate the joint and workspace trajectories to relax the computational effort required. The steepest descent method was applied to obtain a solution, but this approach can, of course, lead to local minimums. However, considerable energy savings were demonstrated in the simulations and experiments with a redundant three-axis planar robot arm. The method is not directly applicable to hydraulic manipulators because the energy objective requires a tailored formulation. The possible solution locality is another concern. A reasonable way to examine solution locality would be to try different initial guesses of the control policy, but it could be time-consuming. Compared with the global DP method, this method may be computationally cheaper but probably sub-optimal.

Regarding DP, a joint-limited solution for general non-redundant manipulators was implemented by Balkan [84]. Minimisation of task execution time and actuator energy consumption, or a hybrid minimisation of the two, was studied. Simulations with a 6-DOF robot arm demonstrated the method's efficacy.

Shin [85] also proposed a DP method for general non-redundant manipulators, which used the phase plane to reduce the effect of the curse of dimensionality on computational complexity. The desired joint trajectories were time-independently parameterised, and the cost functions were expressed in the phase plane. This paper analysed the convergence of the solution and presented interesting numerical examples with electrical manipulators.

Furthermore, Field [86] described an iterative DP approach for minimum-energy trajectory planning for robotic manipulators. This phase-space approach accounted for the manipulator dynamics and proposed a uniform cubic B spline parametrisation of the joint trajectories. Because the DP approach can be computationally expensive, the author proposed the iterative dynamic programming method. Although iterative DP is not really a global algorithm, the experiments still demonstrated significant energy savings with an industrial manipulator. One observation is that the algorithm should be initialised properly so that the solution obtained is optimal. In the methods described above, redundancy resolution at a hydraulic level was not discussed.

Minimum-Time Solutions

The articles in the category of optimal control solutions which have been designed for non-redundant hydraulic manipulators include the article by Enes [87]. The authors presented the conditions that define the hydraulic valve controls when the objective is minimum-time Cartesian control. The authors also derived the conditions under the assumption that the pump-flow rate is limited.

Holobut [88] proposed another minimum-time solution for non-redundant hydraulic excavators. The execution time was parameterised as a monotonically increasing dimensionless function, and the joint trajectories were expressed as a function of the desired end-effector path. However, the author acknowledged the locality of the solution.

The minimum-time solutions designed for redundant hydraulic manipulators include the solutions by Löfgren [8], Ortiz Morales [82] and Mettin [89]. Löfgren [8], for example, presented a global minimum-time solution using the dynamic programming method. The typical 4-DOF hydraulic manipulators were covered. The author also described the maximisation of the lifting capacity of hydraulic manipulators to increase work productivity [90]. Ortiz Morales [82] has addressed semi-automation of typical 4-DOF forestry manipulators. For this purpose, time-optimal joint trajectories were chosen in the phase-space. Experimental results with a heavy-duty commercial and laboratory hydraulic manipulator validated the time-efficiency of the joint trajectories. Mettin [89] presented a similar minimum-time path-constrained joint trajectory generator for forestry manipulators. Mettin's time-optimal solutions can be found in the thesis [91]. In that thesis, using the total mechanical work as a cost function in the framework is briefly discussed. This allows for a time-independent cost function.

Another time-optimal solution is the solution proposed by Ma [92]. This approach enforces joint torque limits, but the important joint velocity limits are not treated. For this reason, it is unsuitable for hydraulic manipulators. Other minimum-time solutions for general manipulators include the methods proposed by Shin [93], Bobrow [94] and Pfeiffer [95], that did consider joint velocity limits.

Regarding non-redundant general manipulators, Constantinescu [96] presented an interesting method, which yields smooth minimum-time joint trajectories subject to actuator

torque and actuator torque rate limits. These actuator torque constraints were transformed into joint jerk, acceleration and velocity limits. The author solved the optimal control problem by using an unconventional flexible tolerance method. Because of the inconvenience related to imposing the joint limits, this method is not suitable for hydraulic manipulators.

In the end, minimum-time solutions always aim to maximise work productivity. These solutions, in general, achieve this by requiring high energy consumption by the system. Minimum-time solutions may also decrease the manipulator's lifetime by straining the components. Energy-optimal redundancy resolutions are better for these issues.

Selecting a suitable task execution time to balance energy consumption and work productivity is important. From this perspective, the minimum-time solution should still be the alternative solution which complements the energy-optimal solution. Finally, an excellent introduction to the minimum-energy, minimum-time and minimum-jerk solutions can be found in [97].

Other Solutions

Besides the minimum-time or minimum-energy solutions, minimum-joint torque schemes have been proposed. Choi [98], for example, minimised joint torque disturbances, i.e., non-constant gravity, inertia, Coriolis and centrifugal terms. The optimal control problem was solved by using DP and kinematic redundancy to reduce system dimensionality. The scheme proposed was said to increase the performance of conventional controllers which are non-robust to load disturbances. The application was to a three-link planar manipulator with rotational joints. Comparing the solution with another joint torque scheme demonstrated that the objective is feasible. Hydraulic manipulators and joint limits, however, were not handled. There are also a myriad of other optimal control solutions and redundancy resolutions. These solutions were omitted.

Summary

Kinematic redundancy is a complex property. The manipulators which have this property have many inverse kinematics solutions. It is, however, a technically convenient property, that can be utilised to optimise joint velocity, energy costs etc. To summarise the main findings, we found virtually no suitable energy-optimal redundancy solutions for hydraulic manipulators considered. Many of the sub-optimal solutions are optimal with electric systems, but they may not be optimal from the standpoint of the hydraulic system. Furthermore, many of the solutions do not consider the joint limits which affect hydraulic manipulators. A number of minimum-time solutions for hydraulic manipulators, on the other hand, were found; see [82, 89]. Judging based on these main findings, the globally energy-optimal redundancy resolutions for hydraulic manipulators are required.

3.3 Numerical Solutions to Optimal Control Problems

There are many numerical methods for solving optimal control problems. A short introduction to the methods, hence, is necessary. But before going into the general properties of these numerical methods, we should emphasise that the pseudo-inverses and

other local approaches are solutions which minimise the cost instantaneously. The optimal control solutions are, in turn, global solutions. Local problems, especially pseudo-inverses, can be solved using numerical integration. Optimal control problems, on the other hand, are commonly solved by employing a specialised optimisation routine.

The numerical methods for solving optimal control problems can be categorised into the following [10, 99]:

1. indirect methods,
2. direct methods and
3. dynamic programming.

Indirect Methods

Indirect methods are of the so-called type optimise first, then discretise [100]. These methods are based on the fundamental theory of optimal control, Pontryagin's minimum principle, which provides the necessary optimality conditions. Pontryagin's minimum principle is founded on minimisation of the Hamiltonian, which conveniently collects the objectives and system dynamics into a multivariate function. By following the guidelines of the established optimal control theory, minimisation of the Hamiltonian leads to a system of ordinary differential equations, from which by solving a two-point or multipoint boundary value problem we obtain the optimum control and corresponding state trajectories. Optimal boundary value problems are usually discretised to obtain the numerical solutions. Overall, Pontryagin's minimum principle is a means of obtaining the equations through which the optimal control problem can be solved.

Because the indirect Pontryagin principle provides only necessary and not sufficient conditions, minimisation or even the existence of the optimal control is not certain. In some problem settings, though, Pontryagin's minimum principle yields boundary value problems simple enough to have an analytical solution. However, numerical methods, for instance, a gradient-based steepest descent, multiple shooting or collocation method, have to be employed in most cases, particularly with regards to nonlinear systems, to solve these boundary value problems. Still, the difficulty of solving differential equations subject to boundary conditions originates from the aforementioned two-point or multipoint boundary value problem (see [101]). Interestingly, Pontryagin's minimum principle is connected to the calculation of variations (variational method) via the Euler-Lagrange equation, which is famously known for providing path solutions that minimise a time-integral of a function dependent on the path variable itself and its time derivative [72]. A reputable and short mathematical introduction to Pontryagin's minimum principle and optimal control theory can be found in Todorov [102].

As for redundancy resolutions and indirect methods, Nakamura [103], Martin [104] and Kim [105] have developed the necessary conditions that yield the optimal joint trajectories which minimise a general integral objective depending on the joint positions and velocities. More recently, Callies [106] discussed a new approach to automate optimal control problem generation and the accompanying boundary value problem through a recursive framework. Nonlinear constraints on the control and states were considered. A multiple-point shooting method was used in the numerical solution of the boundary value problem.

The indirect methods are generally suitable only for "easy" problems. Adding to the fact that indirect methods may not be practical for complex problems, enforcing state constraints, specifically, may be very complicated. Callies' methodology [106], however, simplifies the formulation and yields the redundancy resolution for rigid-link manipulators. Still, Pontryagin's minimum principle may lead only to local optimality if the problem is non-convex. Some problems, though, such as the linear quadratic problem, reduce to an initial value problem under Pontryagin's theory. Application of Pontryagin's minimum principle to these simpler problems is well justified.

Many of the boundary value problems that arise from the necessary conditions of Pontryagin's minimum principle can be solved in Matlab by using the routine bvp4c. This is a collocation method. A tutorial by Wang [107] discusses the Matlab solution of optimal control problem using indirect methods. The tutorial, specifically, describes the use of the Symbolic Toolbox to formulate optimal control problems. The steepest descent method is used for the problems, and the conditions under which bvp4c is effective are discussed.

Direct Methods

Direct methods are of the so-called type discretise first, then optimise [100]. Discretisation of the optimal control problem converts an infinite-dimensional problem involving system dynamics, time and constraints in a finite, constrained nonlinear problem. The multiple shooting method is often utilised in the conversion. The states and controls are approximated using function approximations, and they are evaluated at discrete time stages. The controls are, for example, defined as piecewise constant over the time stages. The state dynamics is discretised using conventional numerical integration methods, and the dynamics are imposed by setting equality ("defect") constraints for all of the time stages. Finally, the states and controls, which are evaluated at the time stages, are set as the optimisation variables. This discretisation procedure enables the application of nonlinear programming problem (NLP) solvers [108].

Direct methods are better suited for complex problems than indirect methods because direct methods enable easier software solution of the optimal control problem. Other advantages include the availability of NLP solvers and the lack of need for analytic differentiations. The major drawback of direct methods is their susceptibility to convergence issues with some constraints and in the presence of discontinuities. A disadvantage concerning both methods is the complexity of globally solving a non-convex, multi-minima minimisation problem. Indeed, depending on the problem and the numerical solver used, a local minimum may be found. This is obviously an unfortunate drawback. In essence, the method of choice for a particular problem depends chiefly on the objective and problem properties, such as its complexity, constraints, discontinuities and the desired scope of the solution. Numerical direct optimisation methods are now widely used in applications [99, 106, 108].

The direct methods implemented in Matlab's native toolboxes are well-suited for solving small-scale optimal control problems. Becerra [109] illustrated the use of such Matlab tools, including the Simulink solvers and nonlinear programming algorithms, such as the fmincon routine, for system simulation and for solving constrained optimal control problems, respectively. The Matlab NLP routines have implementations of the interior-point and sequential quadratic algorithms. These algorithms proceed iteratively from an initial guess towards the optimal policy, but the solution's true optimality is influenced by the algorithm's convergence parameters and the initial guess. NLP solvers other

than the Matlab tools include Snopt [110], Knitro [111] and Ipopt [112]. Amongst the general-purpose optimal control software packages are Tomlab (Propt) [113], GPOPS-II [114], DIDO [11] and ACADO Toolkit [115]. These packages are available for Matlab, but only the ACADO Toolkit is free. The DIDO software claims to be the only solver which does not require an initial guess of the optimal control. Many of the contemporary optimal control solvers, like the DIDO solver, are based on pseudospectral optimal control theory.

Dynamic Programming

Dynamic programming [10, 116, 117] is based on Richard Bellman's famous principle of optimality, which can be used to solve optimal control problems to global optimality. Conventional direct optimisation methods, by comparison, can produce local results in some cases. Particularly for this reason, the DP algorithm was chosen to solve the optimal control problems presented in this research work.

The DP algorithm is significantly more effective than a brute-force method, whose computational cost grows exponentially with the time stages N. This effectiveness arises from the principle of optimality. For demonstration purposes, let us present the DP algorithm for the scalar case. In this case, the principle of optimality can be stated in discrete-time using the so-called Bellman equation as in

$$J_k^\star\big(x_k\big) = \min_{u_k \in \mathbb{U}} \Big\{ L_k\big(x_k, u_k\big) + J_{k+1}^\star\big(F_k(x_k, u_k)\big) \Big\} \tag{3.15}$$

where k denotes the discrete time stage, x_k and u_k are the state and controls, respectively, L_k is the running cost to be minimised over the time horizon and F_k denotes the discrete system dynamics. If we consider that the optimisation time is divided into N intervals of equivalent length, then the minimum cost-to-go $J_k^\star$ is the minimum sum of the running costs from stage $k \in \{0, 1, \ldots, N-1\}$ to the final stage at $k = N-1$. We define this minimum cost-to-go $J_k^\star\big(x_k\big)$ for each state vector combination such that $x_k \in \mathbb{X}$, and we minimise it over all the possible controls at stage k such that $u_k \in \mathbb{U}$. Here, $\mathbb{X} = \{x^{(1)}, x^{(2)}, \ldots, x^{(N_\mathrm{x})}\}$ and $\mathbb{U} = \{u^{(1)}, u^{(2)}, \ldots, u^{(N_\mathrm{u})}\}$ denote the discrete state and control sets, respectively. The variables N_x and N_u, in turn, determine the number of discretised state and control values, respectively, in the "grids". Sufficiently dense grids and time discretisation must be defined to reach the global optimum with continuous-time systems whilst considering that the computational and memory requirements of the algorithm scale linearly with the density of the grids.

However, this algorithm is mainly suited for the solution of lower-dimensional problems because of its computational inefficiencies. Indeed, the algorithm's main disadvantage is that its computational complexity grows exponentially with the states and controls due to the curse of dimensionality. But when the problem size is limited to several states and controls, DP is a compelling alternative to consider in the solution of difficult optimal control problems with discontinuities and complexities in general. These types of difficult optimal control problems can arise when dealing with problems related to hydraulic systems because of the inherent properties of these systems. Note that actuator limits can be enforced in the algorithm by adding a significant penalty to the running cost when the limits are not obeyed.

Algorithm 3.3.1 Discrete-time dynamic programming algorithm.

1. Backward phase with interpolation

Require: Sets $\mathbb{X}$ and $\mathbb{U}$

Initialisation: set final cost-to-go $J_N^\star\big(x_N\big)$ for all $x_N \in \mathbb{X}$
for $k = N-1$ to 0 **do**
    for all $x_k \in \mathbb{X}$ **do**
        for all $u_k \in \mathbb{U}$ **do**
            $x_{k+1} = F_k(x_k, u_k)$
            Interpolate $J_{k+1}^\star(x_{k+1})$
            $J_k\big(x_k, u_k\big) = L_k(x_k, u_k) + J_{k+1}^\star(x_{k+1})$
        end for
        $J_k^\star\big(x_k\big) = \min_{u_k \in \mathbb{U}} J_k\big(x_k, u_k\big)$
        $u_k^\star\big(x_k\big) = \operatorname{argmin}_{u_k \in \mathbb{U}} J_k\big(x_k, u_k\big)$
    end for
end for

2. Forward simulation with interpolation

Require: $u_k^\star(x_k)$ for all $k \in \{0, 1, \ldots, N-1\}$ and $x_k \in \mathbb{X}$

Initialisation: set x_0
for $k = 0$ to $N-1$ **do**
    Interpolate $u_k^\star(x_k)$
    $x_{k+1} = F_k\big(x_k, u_k^\star(x_k)\big)$
end for

The general DP algorithm is presented in Algorithm 3.3.1. The algorithm comprises a backwards phase and a forwards simulation phase. In the backwards phase, the optimal controls are solved backwards in time by applying the Bellman equation. Interpolation of the cost-to-go is usually required at this stage to compute the next-state minimum running costs. Note that we have denoted the optimal control with $u_k^\star\big(x_k\big)$ in the algorithm. In the forwards phase, the system is simulated forwards in time using the optimal controls computed in the backwards phase. At this stage, interpolation of the optimal controls at the current state is also required. When the algorithm finishes, it has produced the optimal control sequence and the system trajectory.

To summarise, although the DP algorithm can be computationally costly if the problem size is too large, the computational complexity is inherently reduced compared to a brute-force algorithm through clever division of the larger problem into simpler sub-problems via the principle of optimality/Bellman equation. The DP method is discussed in more detail together with the redundancy resolutions proposed [19, 20].

4 | Summary of Publications

This thesis as a compendium comprises *six* research publications, each of which is briefly summarised below. Publications **P-I**–**P-IV** deal with condition monitoring of hydraulic manipulators, and publications **P-V** and **P-VI** discuss energy-optimised redundancy resolution of hydraulic manipulators from the perspective of valve-controlled hydraulic systems.

4.1 P-I: Detection and Isolation of Leakage and Valve Faults in Hydraulic Systems in Varying Loading Conditions, Part 1: Global Sensitivity Analysis

This paper introduces a systematic approach based on GSA for the determination of the condition monitoring model's sensitive parameters. The studied model parameters include the nominal flow rate and the nominal pressure difference at the pressure and return notch of the hydraulic proportional valve, effective bulk modulus, spool offset of the proportional valve and viscous friction coefficient. These parameters are varied uniformly within selected limits, and the parameters' effect on the pressure response, including cross-effects, is analysed using first-order and total-order sensitivity indices. Moreover, the pressure effects of cylinder leakages are analysed by utilising GSA.

The main contribution of this paper is the results demonstrating that the flow coefficient of the supply notch is the most important parameter to accurately verify in the condition monitoring model, followed by the flow coefficient of the return notch. The sensitivity of the bulk modulus is also significant, whereas the influence of friction is negligible. The GSA applied to analyse the pressure effects of cylinder leakages demonstrates that external and internal leakages of the cylinder actuator, or alternatively, the hydraulic valve, can be pressure-identified based on the sensitivity of these leakages on the pressure responses.

4.2 P-II: Detection and Isolation of Leakage and Valve Faults in Hydraulic Systems in Varying Loading Condition, Part 2: Fault Detection and Isolation Scheme

In this follow-up paper, a model-based FDI scheme for multi-DOF hydraulic manipulators controlled using hydraulic proportional valves is proposed. Exploiting our GSA results, the valve flow coefficients are meticulously identified at steady-states to increase the robustness

of the model-based condition monitoring solution. This proposed condition monitoring solution is UKF-based, and it utilises experimentally verified adaptive thresholds and moving average-filtered pressure residuals to increase the robustness of the method towards minor modelling errors. Moreover, the approach is load-independent because the unknown load dynamics are compensated by using a joint position measurement.

The scheme is experimentally verified on a test-bed, which was equipped with needle valves to emulate internal and external leakages. Furthermore, simulations of valve faults are performed using a verified model of the test-bed. Fault residual patterns are verified to enable the isolation of

- external leakage from the piston-side,
- external leakage from the rod-side,
- internal leakage and
- spool jamming to different positions (spool offset).

Experimental results demonstrated that an external leakage of class 0.20 L/min was identifiable. The minimum identifiable internal leakage level was not sought.

The main contribution of this paper is the FDI scheme designed for hydraulic subsystems consisting of hydraulic cylinder actuators controlled by hydraulic proportional valves. Fault detection in the condition monitoring solution is achieved without knowledge about the load dynamics on the hydraulic cylinder actuator. Therefore, this proposed scheme is highly applicable to hydraulic manipulators which are controlled by using proportional valves because the unknown and potentially rapidly changing load in these systems poses no real threat to fault detection robustness. However, fault isolation, is achieved based on extended residual patterns. The peak fault detection sensitivity achieved in this publication is amongst the state-of-the-art in condition monitoring of hydraulic systems. However, the fault detection sensitivity of the proposed condition monitoring solution was time-variant, as was demonstrated through experiments.

4.3 P-III: Micro-Electromechanical System Sensors in Unscented Kalman Filter-Based Condition Monitoring of Hydraulic Systems

The industry interest in introducing intelligent sensor-based functionality to hydraulic manipulators has led to the introduction of low-cost MEMS-based motion sensors. These rugged, easy-to-install sensors incorporate sophisticated signal processing methods, and the sensors can be employed in a geometry-aided approach to produce an accurate, low-noise and lag-free full motion-state for a rigid-body manipulator [118, 119]. Compared to the difficult installation of in-cylinder position sensors, MEMS sensors require no physical contact with the actuators or joints. The 10 centimetres required by the in-cylinder sensors is about a fifth of the typical cylinder stroke, and it amounts to a significant positional change at the end-effector of the hydraulic manipulator. Hence, MEMS sensors could be suitable for hydraulic manipulators.

This paper analyses the suitability of MEMS sensors to the condition monitoring task designed in **P-II**. The author was the principal writer of this publication, but Drs.

Janne Honkakorpi from IHA (TUT) and Juho Vihonen from SGN (TUT) contributed, particularly, to the writing of Section II and suggested improvements to the paper in general.

The main contribution of this publication is the experimental validation of the UKF-based FDI method by exploiting a low-cost MEMS motion sensor. The results demonstrate that the MEMS sensor provides a sufficiently accurate motion-state for the condition monitoring task. The suitability of the low-cost MEMS motion sensor to the condition monitoring task represents a step towards intelligent machines which are augmented with sensor-based software solutions.

4.4 P-IV: Detection and Isolation of Faults in Mobile Hydraulic Valves Based on a Reduced-Order Model and Adaptive Thresholds

This paper introduces a parity space-based FDI scheme for multi-DOF hydraulic systems controlled using mobile hydraulic valves. Due to the hydraulic valve's complexity (two-stage pilot-operation, dead-zone and pressure-compensation), a sufficiently simple hydraulic model is proposed. This model is verified to enable fault subset detection. A procedure to statistically determine adaptive residual thresholds is designed to increase robustness to false and missed alarms. Fault residual patterns are verified to enable the isolation of the following faults:

- external leakage from the piston-side,
- external leakage from the rod-side,
- internal leakage,
- velocity sensor bias,
- cylinder pressure sensor biases,
- compensator pressure sensor bias,
- pilot pressure sensor biases,
- main spool locking,
- main spool erosion,
- compensator spool locking,
- compensator spool erosion and
- pilot pressure spool offset.

Experiments on a commercial hydraulic manipulator are performed to validate the fault detection sensitivity of the designed scheme. These experiments demonstrate that additive pressure sensor biases as minor as $0.3-0.5$ MPa might be detectable in some cases. However, the fault sensitivity of the proposed method varied because of modelling errors.

The main contribution of this paper is the FDI scheme designed for hydraulic subsystems comprising hydraulic cylinder actuators controlled by mobile hydraulic valves. Fault detection is independent of load dynamics, as in the publications **P-II** and **P-III**. In addition, the robustness of the method to modelling errors is enhanced by employing statistically verified adaptive thresholds and a simplified hydraulic model. The downside of this scheme is the possibility of detecting a subset of the faults in the full fault table. This is a consequence of the additional hydraulic pressure measurements required to increase the reliability of fault detection.

4.5 P-V: Global Energy-Optimised Redundancy Resolution in Hydraulic Manipulators Using Dynamic Programming

This paper proposes a DP-based hydraulic energy-optimal solution to the open kinematic redundancy problem of planar 3-DOF hydraulic manipulators. Separate metre-in and separate metre-out systems were particularly considered in this study.

The main contributions of this paper are the optimal control problem formulation and the hydraulic energy-optimal kinematic redundancy resolution presented. Our formulation includes the design of cost functions that approximate the hydraulic energy use of the CP and LS systems. The formulation also includes the proposed modelling approach, in which only the redundant extension cylinder motion is optimised. The motion-states of the non-redundant joints, instead, are resolved using inverse kinematics. The solution proposed also satisfies a prescribed, twice time-differentiable Cartesian path and actuator limits (strokes, velocity and acceleration).

The solution proposed resolves the kinematic redundancy more effectively from the perspective of hydraulic energy minimisation than do point-wise optimal pseudo-inverses or standard optimal control tools. The simulation results motivate automation and facilitate joint trajectory optimisation in these future applications at construction sites.

4.6 P-VI: Global Energy-Optimal Redundancy Resolution of Hydraulic Manipulators: Experimental Results for a Forestry Manipulator

This paper extends the energy-optimised redundancy resolution of our previous paper to non-planar 4-DOF hydraulic manipulators to solve the open problem concerning energy-optimised redundancy resolution of 4-DOF hydraulic manipulators. Experimental results are presented to demonstrate the high performance of extended energy-optimal redundancy resolution for the first time. A commercial forestry hydraulic manipulator was equipped with LS and CP hydraulic systems for experimental evaluation. Around 15 – 30% higher energy consumption was observed in the LS and CP systems with conventional methods compared to the proposed hydraulic energy-optimal redundancy resolution. For energy-optimation purposes, a practical mathematical model of the complex experimental test-bed was systematically presented and verified in this study.

The main contributions of this paper include the experimental performance evaluation of the proposed redundancy resolution and our practical minimum-state mathematical model. For example, kinematic redundancy was used to reduce the optimal control problem size

significantly from eight states and four controls to two states and one control. Nevertheless, manipulator dynamics, pressure-compensation including the cylinder back-pressures of the two most common spool configurations of mobile hydraulic valves, were considered in the proposed system model. Moreover, a prescribed end-effector path and position, velocity and acceleration bounds of cylinder actuators were satisfied in the solution.

The experimental performance evaluation of the proposed redundancy resolution suggests that minimisation of the pump flow rate, that minimises the CP system's energy consumption, often leads to a satisfactory hydraulic energy consumption irrespective of the hydraulic system (CP or LS). Moreover, this pump flow rate minimisation is simpler to formulate than minimising the energy consumption of actuators or the LS system. As in the previous paper, these energy-optimal joint trajectories can be saved to a database, from which they can be retrieved when needed for real-time use.

5 | Discussion

This section presents a summary of how the RP, divided into RP1 – RP6 as described in Section 1.1, is addressed.

5.1 Load-Independency (RP1)

The condition monitoring solution should be load-independent because several valve-controlled hydraulic cylinders on heavy-duty manipulators are usually subjected to varying load conditions. Thus, the solutions should not require knowledge about the magnitude and variance of the payload disturbance at the manipulator end-effector, which is observed at the hydraulic cylinders.

All the model-based condition monitoring solutions proposed in the research publications **P-II** – **P-IV** [16 – 18] are load-independent. This is a significant achievement because of the considered application to heavy-duty hydraulic manipulators where the external forces on the hydraulic cylinders vary unpredictably. As is discussed in **P-II**, load-independent condition monitoring solutions are derived by conveniently replacing the motion dynamics of the hydraulic cylinders with joint sensor measurements (see also [45, 52]). Based on these joint measurements, the cylinder positions and the cylinder velocities can be calculated at an accuracy comparable to that of a highly accurate system model. Because modern hydraulic manipulators are more and more often equipped with motion sensors [2 – 4], achieving load-independency of model-based condition monitoring by utilising joint position measurements is possible.

5.2 General Applicability (RP2)

The condition monitoring solution should be applicable to multi-actuator hydraulic manipulators which are valve-controlled using proportional hydraulic valves or mobile hydraulic valves with complex built-in hydro-mechanical functionalities.

Because the hydraulic actuators of hydraulic manipulators are predominantly valve-controlled, it was reasonable to limit the thesis's scope and examine only valve-controlled hydraulic manipulators. However, because the actuators of hydraulic manipulators can be valve-controlled by using either proportional hydraulic valves or complex pilot-operated, pressure-compensated mobile hydraulic valves, condition monitoring solutions were required to be developed for both valve types.

Firstly, the UKF-based condition monitoring solution designed in **P-II** is applicable to multi-actuator hydraulic manipulators which are controlled using proportional hydraulic

valves. This model-based solution is load-independent, as has been discussed. Therefore, the application of this condition monitoring solution requires knowledge only about the hydraulic system parameters. Valve parameters, for example, were meticulously identified for the system model based on the results of the GSA presented in **P-I** [15]. The model-based condition monitoring solution designed in **P-II** was extended in **P-III** to hydraulic manipulators which use a low-cost MEMS motion sensor. This extension experimentally verified that low-cost MEMS motion sensors can be used in the proposed condition monitoring solution instead of high-cost reference joint encoders, which is significant from the perspective of the solution's general applicability.

Secondly, a model-based condition monitoring solution for multi-actuator hydraulic manipulators controlled using mobile hydraulic valves with complex built-in hydro-mechanical functionalities was designed in **P-IV**. This condition monitoring solution is also load-independent to facilitate its application to heavy-duty hydraulic manipulators. However, this model-based solution designed in **P-IV** might have limited applicability to hydraulic manipulators where all the necessary hydraulic pressure measurements are not available. The measurements of hydraulic pilot pressures at the ends of the main spool, instead of the valve control current measurement used in the proposed proportional valve solutions, and the measurement of hydraulic "supply pressure" upstream of the main spool, instead of the supply pressure measurement of the hydraulic pump, are necessary for this condition monitoring solution designed for mobile hydraulic valves.

5.3 High Performance (RP3)

The condition monitoring solution should be fault-sensitive and capable of isolating common hydraulic system faults. The FDI performance should be experimentally evaluated.

High performance was expected from the condition monitoring solutions in FDI. For this reason, model-based condition monitoring solutions which can achieve better fault isolation performance compared with non-model-based methods were designed. A foundation for high performance in FDI was laid in **P-I** with the GSA results that determined the most sensitive system parameters for tuning of the condition monitoring model. Publications **P-II**–**P-IV** built on the foundation laid in **P-I** by concentrating on the identification of the sensitive model parameters, particularly the hydraulic valves' flow coefficients.

High performance in fault detection was experimentally demonstrated in **P-II** and **P-III** with external and internal leakages. For example, detecting an external leakage as small as 0.17 L/min was possible in some cases. In addition, the method proposed for mobile hydraulic valves in **P-IV** detected sensor biases smaller than 0.5 MPa in the experiments. These results are state-of-the-art. However, fault detection sensitivity varied as a function of time due to the modelling errors observed in the condition monitoring models of the proposed solutions.

As for fault isolation performance, common internal leakages, external leakages, jamming faults of the valve spool and positive sensor biases were isolated correctly with high accuracy in the experiments presented in **P-II**–**P-IV**. However, valve jamming faults could not always be distinguished from internal leakages because of their similar residual pattern. In addition, new fault isolation patterns were presented for the mobile hydraulic valve and the cylinder actuator subsystem in **P-IV**. These new fault patterns were identified through simulations. Achieving consistently high FDI performance was, however,

difficult because of the modelling errors affecting the system models of the condition monitoring solutions.

5.4 Energy-Optimality (RP4)

Kinematic redundancy resolution should be energy-optimal at the valve-controlled hydraulic actuator and hydraulic power system interaction level. The minimum attainable energy consumption of the commonly-used LS and CP hydraulic systems should be obtained.

The kinematic redundancy resolutions proposed for hydraulic manipulators in **P-V** [19] and in **P-VI** [20] are energy-optimal at the valve-controlled hydraulic actuator and hydraulic power system interaction level. Hydraulic energy consumption, in particular, was minimised because system energy efficiency is a major problem in valve-controlled hydraulic manipulators [6]. Furthermore, we considered the minimisation of hydraulic energy consumption instead of actuator energy consumption because the energy consumption of the hydraulic system does not equate to the energy consumption of the actuators due to the pressure losses encountered in valve-controlled hydraulic systems. The sub-optimality of the strategy minimising the energy consumption of the actuators and the sub-optimality of pseudo-inverses in hydraulic energy minimisation were demonstrated through simulations in **P-V** and through experiments in **P-VI**.

The most significant contribution of this thesis is this energy-optimal kinematic redundancy resolution of hydraulic manipulators, including the formulation of optimal control problems whose global solution guarantees the energy-optimality of redundancy resolution at the hydraulic system level. Optimal control problems with effective hydraulic system cost-functions were formulated specifically for the commonly used LS and CP systems. Moreover, the optimal control problem size was reduced to the bare essentials so that the global solution could be obtained using DP. To reduce the problem size, thus computational complexity, firstly, only the motion of the redundant extension cylinder was optimised without causing any apparent tracking error in the desired end-effector path, for example. Secondly, steady-state hydraulic equations were derived for the typical pressure-compensated, non-differential spool valves and the pressure-compensated, differential spool valves and were incorporated in the proposed cost-functions to obtain the desirable effective solution.

5.5 General Applicability (RP5)

Kinematic redundancy resolution should be applicable to typical kinematically-redundant 3-DOF and 4-DOF hydraulic manipulator designs.

To achieve the broadest applicability for the proposed hydraulic energy-optimised redundancy resolution, this redundancy resolution was designed for common 3-DOF and 4-DOF hydraulic manipulators in **P-V** and **P-VI**, respectively. Indeed, most of the heavy-duty hydraulic manipulators have a similar 3-DOF or 4-DOF kinematically redundant serial-link configuration, which is typically either of the RRP-type for the vertical xy-plane or the RRRP-type for the xyz-space. The R stands for the revolute joint and the P for the prismatic joint in this standard naming scheme. In other words, hydraulic manipulators typically have two to three hydraulically actuated revolute joints followed

by a hydraulically-actuated prismatic joint to extend the reach of the manipulator, which also makes these hydraulic manipulators kinematically redundant by 1-DOF.

This kinematic redundancy was put to effective use in the designed redundancy resolutions in **P-V** and **P-VI** by optimising only the motion of the redundant joint with respect to minimisation of hydraulic energy consumption. The motions of non-redundant joints can be readily derived from the optimised motion of the redundant joint by employing kinematic inversion equations at the desired end-effector path.

5.6 High Performance (RP6)

Kinematic redundancy resolution should provide high performance whilst satisfying the position, velocity and acceleration bounds of the actuators. The performance of the redundancy resolution should be experimentally assessed by comparison to conventional methods.

High-performance kinematic redundancy resolution was one of the main objectives in this thesis. This requirement for high performance entailed that the position, velocity and acceleration limits of the actuators are satisfied in the redundancy resolution whilst the energy consumption of the valve-controlled hydraulic system (CP or LS) is minimised. Furthermore, the hydraulic manipulator's end-effector had to track a desired path.

High performance of the proposed redundancy resolution was observed in the simulations presented for a 3-DOF construction crane model in **P-V** and the experiments presented for a complex 4-DOF forestry hydraulic manipulator in **P-VI**. Around 15 – 30% greater energy consumption was observed with the conventional methods compared to the proposed redundancy resolution in the LS and CP hydraulic systems in both studies. In addition, minimising the energy consumption of actuators was sub-optimal in some end-effector paths. These significant results were obtained whilst the cylinder actuators of the hydraulic manipulators satisfied the actuator position, velocity and acceleration limits and the manipulator end-effector traced the desirable path.

Because comparison of the proposed methods to widely adopted methods is one of the cornerstones of academic research, a representative comparison of conventional methods combined with suitable minimised objectives against the proposed solutions was provided to study the solutions' effect on the energy consumption of the valve-controlled LS and CP systems through simulations in **P-V** and through experiments in **P-VI**.

6 | Conclusions and Future Work

This thesis covers and proposes FDI and energy-optimal redundancy resolution at the hydraulic system level for hydraulic manipulators on mobile machines. These designed software-based solutions increase the automation level of hydraulic manipulators and are, in some form, expected to have an essential role in future intelligent hydraulic manipulators, particularly with regard to robotic control of hydraulic manipulators.

6.1 Fault Detection and Isolation of Hydraulic Manipulators

Hydraulic manipulator OEMs have often considered condition monitoring an expensive feature which requires advanced computer technologies and a major sensor presence in the system. Hence, sophisticated condition monitoring systems have not been adopted yet. This situation might be slowly changing because manipulator OEMs are becoming increasingly more interested in increasing the automation level of their machines. Furthermore, sensor and computer technologies have reached a point where the introduction of condition monitoring for hydraulic manipulators is more realistic than ever.

FDI is a vast research field, in which a number of model-based and non-model-based methods have been proposed. These model-based methods are particularly suited for fault recognition and fault cause identification. However, not all of these model-based methods are suitable for monitoring hydraulic components in heavy-duty manipulators. Unfortunately, many of these methods had also been studied using computer simulations, which are not affected by realistic modelling errors. Consequently, novel and experimentally-verified solutions were developed for monitoring heavy-duty manipulators in this thesis. The thesis results demonstrate that fluid leakages, both external and internal, as well as the intermittent valve offsets, can be effectively detected and isolated from the other faults by using the model-based condition monitoring schemes designed. However, fault sensitivity is still a challenge because of the modelling errors, which arise from the complexity of mobile hydraulic systems, and because a limited amount of data from the faults and wear effects of hydraulic systems are generally available.

6.2 Redundancy Resolution for Robotic Control of Hydraulic Manipulators

There is strong indication from a number of areas that robotic control is a major trend. Following the example of the automotive industry, for example, hydraulic manipulator

OEMs are looking into semi-automation, in which routine tasks and end-effector movements are transferred from the responsibility of human operators to computer algorithms to improve work productivity, operator workload circumstances and machine operation in general. This semi-automated control approach would obviously be an important step towards autonomous systems.

Energy-efficiency is another soaring trend for heavy-duty machinery. Increasing the energy-efficiency of heavy-duty machinery obviously aims to reduce fossil fuel consumption and the diesel emissions of the machinery. The recently strong movement in this space of energy-efficiency is encouraging because the hydraulic machinery industry is generally quite conservative, and change is often slow.

In accordance with these trends, energy-optimised redundancy resolution for robotic control of hydraulic manipulators is considered. Remarkably, the energy consumption of hydraulic manipulators can be decreased through energy-optimised redundancy resolution, because the typically kinematically-redundant hydraulic manipulators lend themselves to joint motion optimisation. By employing redundancy in an optimised way to find the joint trajectories of least hydraulic energy consumption along a prescribed workspace path, the results of this thesis have demonstrated that the energy consumption of hydraulic manipulators can be reduced, even with industry-standard hydraulic systems. This result is remarkable and encouraging for hydraulic manipulators.

6.3 Connection Between Condition Monitoring and Robotic Control

The trend towards robotic control of hydraulic manipulators ties in with condition monitoring of these machines. Specifically, remote condition monitoring systems exploiting modern information technologies could offer the principal means to ensure the proper functioning of automated machines with minimal human supervision. Thus, condition monitoring solutions would be highly important to the practicality of automated operations of autonomous hydraulic manipulators. Remote condition-monitoring, however, is connected to the fourth industrial revolution, i.e., the revolution of the Industrial Internet. With this revolution of the Industrial Internet, hydraulic machines would increasingly be a means to be connected to a centralised server and database, which deals with condition monitoring tasks. The RP of this thesis, consisting of the designing of condition monitoring solutions and energy-optimal redundancy resolution for heavy-duty hydraulic manipulators is thus connected through the long-term goal of robotic control. Both are also obviously software-based functionalities, that increase the automation level of hydraulic manipulators.

6.4 Future Work

Fault Detection and Isolation of Hydraulic Manipulators

Concerning condition monitoring, adaptation to changing environments would have to be more closely considered. For example, ageing-related wear was not addressed in the designed condition monitoring schemes. This wear, particularly outside of laboratory conditions, could affect the model-based solutions by reducing the modelling accuracy

slowly over time. Overcoming the effects of wear, if they are not included as faults, would require reconfiguration of the hydraulic model periodically offline or online based on parameter adaptation. Temperature-dependence of the valve flow coefficients was also not discussed in the condition monitoring schemes. If the manipulator is operated when the hydraulic fluid is cold or the fluid temperature is far from the ideal temperature, then the valve flow coefficients should be adapted to the temperature changes based on the readings from a fluid temperature sensor. This could be achieved using temperature models pre-determined offline or using online, yet slow, parameter adaptation. However, we suspect that the best approach might be to use the condition monitoring system only when the hydraulic fluid is warm enough.

Determining the correct maintenance time or estimation of the remaining useful life was not considered in the condition monitoring schemes. This requires a more statistical approach, for example, logistic regression applied to the magnitude of the residuals. These logistic regression methods could prove useful in future condition monitoring solutions. Successful employment of statistical methods to condition monitoring might, however, depend heavily on the amount of data available from the faulty conditions to train these methods. Acquiring supplementary data to support fault diagnosis is generally difficult. To determine the optimal maintenance time, the data would have to be obtained from multiple machines operating in realistic conditions, whilst considering the effects of strain and wear on the hydraulic components.

Concerning the modelling in the designed condition monitoring schemes, an accurate pressure-compensator dynamics model was not built, because it proved difficult to do, even after compensator parameters, such as the compensator spool mass and the return spring coefficient, were measured. In hindsight, the pressure-compensator dynamics could have been easier to verify if, for identification purposes, the mobile hydraulic valve had been first installed in a simple linear test-bed. Accurate modelling of the manipulator dynamics could have alternatively been helpful to facilitate identification of the pressure-compensator subsystem. The manipulator dynamics were not considered because the condition monitoring methods used were load model-independent. Ultimately though, modelling pressure-compensator dynamics with its minor hydraulic volumes is difficult and requires an extensive simulation effort because of the numerical stiffening effect these compensator models have on the manipulator's simulation model. Moreover, replacing the pilot pressure measurements by using a model of the main spool dynamics was not considered because the pilot pressure measurements were effectively used as an indirect way of measuring main spool position, the knowledge of which is important for modelling accuracy, thus also fault detection accuracy.

FDI of hydraulic pumps, motors and filters, amongst other hydraulic components of heavy-duty manipulators, were excluded from this research work to limit scope. FDI methods can be developed for these hydraulic components in future work.

Redundancy Resolution for Robotic Control of Hydraulic Manipulators

Concerning redundancy resolution, some properties were also left for future work. Limiting motion jerk in the joint trajectories and, perhaps more importantly, dealing with non-prescribed workspace paths were omitted. Motion jerk limitation might be useful because the manipulator's components could wear more quickly if the joint motions optimised

were too jerky. Optimising over non-prescribed workspace paths would have complicated the problem too much as it would have required a simultaneous resolution of optimal joint trajectories and end-effector positions, subjected to initial, final and way points. Thus, it was omitted. Optimisation potential could, however, increase in this case, which makes it an interesting future problem. However, optimisation over a non-prescribed workspace path might require a different approach than that presented in this work. It might also be possible to just relax the end-effector path constraints to increase the energy-saving potential. Finally, real-time application of the energy-optimal redundancy resolution was left for future work.

Bibliography

[1] "HIAB loader," https://www.hiab.com/fi-FI/Alkusivu/Tuotteet/Kuormausnosturit/HIAB/Tuotteet-hiab/hiab-x-hipro-638, 2017, [Online; accessed 04-March-2017].

[2] "John Deere Intelligent Boom Control, in Finnish," https://www.deere.fi/fi_FI/industry/forestry/learn_more/features/ibc.page, 2017, [Online; accessed 10-March-2017].

[3] "Cranab's crane-tip control," http://www.cranab.se/site_specific/uploaded_files/media/2013/05/cran_fc_2013-05-24_en.pdf, 2017, [Online; accessed 10-March-2017].

[4] "HIAB's crane-tip control," https://www.hiab.com/en/global/about-us/newsroom/news/hiab-crane-tip-control/, 2017, [Online; accessed 10-March-2017].

[5] "Ponsse Scorpion's structural improvement," http://www.ponsse.com/products/harvesters/scorpion, 2017, [Online; accessed 10-March-2017].

[6] J. Mattila, J. Koivumaki, D. Caldwell, and C. Semini, "A survey on control of hydraulic robotic manipulators with projection to future trends," *To appear in IEEE/ASME Trans. on Mech.*, 2017.

[7] J. Watton, *Modelling, monitoring and diagnostic techniques for fluid power systems.* London, England: Springer-Verlag, 2007.

[8] B. Löfgren, "Kinematic control of redundant knuckle booms with automatic path following functions," Ph.D. dissertation, Royal Institute of Technology (KTH), Stockholm, 2009.

[9] P. Frank, S. Ding, and T. Marcu, "Model-based fault diagnosis in technical processes," *Transactions of the Institute of Measurement and Control*, vol. 22, no. 1, pp. 57–101, 2000.

[10] D. E. Kirk, *Optimal control theory: An introduction.* Courier Dover Publications, 2012.

[11] "DIDO optimal control software," http://www.elissarglobal.com/, 2015, [Online; accessed 19-July-2015].

[12] F. Flacco, A. De Luca, and O. Khatib, "Motion control of redundant robots under joint constraints: Saturation in the null space," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2012, pp. 285–292.

[13] L. Beiner and J. Mattila, "An improved pseudoinverse solution for redundant hydraulic manipulators," *Robotica*, vol. 17, no. 2, pp. 173–179, 1999.

[14] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global sensitivity analysis: The primer*. John Wiley & Sons, 2008.

[15] J. Nurmi and J. Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 1: global sensitivity analysis," *International Journal of Fluid Power*, vol. 12, no. 3, pp. 41–51, 2011.

[16] J. Nurmi and J. Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 2: fault detection and isolation scheme," *International Journal of Fluid Power*, vol. 13, no. 1, pp. 17–27, 2012.

[17] J. Nurmi, J. Honkakorpi, J. Vihonen, and J. Mattila, "Micro-electromechanical system sensors in unscented Kalman filter-based condition monitoring of hydraulic systems," in *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Wollongong, Australia, July 9-12*, 2013, pp. 354–361.

[18] J. Nurmi and J. Mattila, "Detection and isolation of faults in mobile hydraulic valves based on a reduced-order model and adaptive thresholds," in *Proceedings of the ASME/BATH Symposium on Fluid Power and Motion Control (FPMC), Sarasota, Florida, USA, October 6-9*, 2013, pp. V001T01A020–V001T01A020.

[19] J. Nurmi and J. Mattila, "Global energy-optimised redundancy resolution in hydraulic manipulators using dynamic programming," *Automation in Construction*, vol. 73, pp. 120–134, 2017.

[20] J. Nurmi and J. Mattila, "Global energy-optimal redundancy resolution of hydraulic manipulators: Experimental results for a forestry manipulator," *Energies*, vol. 10, no. 5, p. 31, 2017.

[21] R. Isermann, *Fault-diagnosis systems: An introduction from fault detection to fault tolerance*. Berlin, Germany: Springer-Verlag, 2006.

[22] "MTS Sensors: Mobile hydraulic sensors," http://www.mtssensors.com/products/mobile-hydraulic-sensors, 2015, [Online; accessed 21-June-2015].

[23] R. Patton and J. Chen, "Observer-based fault detection and isolation: Robustness and applications," *Control Engineering Practice*, vol. 5, no. 5, pp. 671–682, 1997.

[24] P. M. Frank and X. Ding, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *Journal of Process Control*, vol. 7, no. 6, pp. 403–424, 1997.

[25] M. Benini, P. Castaldi, and S. Simani, *Fault diagnosis for aircraft system models: An introduction from fault detection to fault tolerance*. VDM Publishing, 2009.

[26] R. J. Patton and J. Chen, "Advances in fault diagnosis using analytical redundancy," in *Proceedings of the IEEE Colloquium on Plant Optimisation for Profit (Integrated Operations Management and Control, Digest No. 1993/019)*, 1993, pp. 6–1.

[27] R. Isermann, "Model-based fault-detection and diagnosis – Status and applications," *Annual Reviews in Control*, vol. 29, no. 1, pp. 71–85, 2005.

[28] M. Witczak, *Modelling and estimation strategies for fault diagnosis of non-linear systems: From analytical to soft computing approaches*. Springer Science & Business Media, 2007, vol. 354.

[29] S. Simani, C. Fantuzzi, and R. J. Patton, *Model-based fault diagnosis in dynamic systems using identification techniques*. Springer Science & Business Media, 2013.

[30] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 636–653, 2010.

[31] R. Isermann and P. Balle, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control Engineering Practice*, vol. 5, no. 5, pp. 709–719, 1997.

[32] H. E. Merritt, *Hydraulic control systems*. John Wiley & Sons, 1967.

[33] M. Jelali and A. Kroll, *Hydraulic servo-systems: Modelling, identification and control*. Springer Science & Business Media, 2003.

[34] R. Isermann and M. Münchhof, "Identification of dynamic systems: An introduction with applications (advanced textbooks in control and signal processing)," 2011.

[35] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Syst. Mag.*, vol. 8, no. 6, pp. 3–11, 1988.

[36] S. X. Ding, *Model-based fault diagnosis techniques: Design schemes, algorithms, and tools*. Springer Science & Business Media, 2008.

[37] J. Gertler, "Analytical redundancy methods in fault detection and isolation," in *Preprints of IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, 1991, pp. 9–21.

[38] D. Simon, *Optimal state estimation*. Hoboken, NJ: Wiley, 2006.

[39] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Springer Science & Business Media, 2012, vol. 3.

[40] R. Isermann, *Fault-diagnosis applications. Model-based condition monitoring: Actuators, drives, machinery, plants, sensors, and fault-tolerant systems*. Springer, 2011.

[41] R. Isermann, *Mechatronic systems: Fundamentals*. Springer Science & Business Media, 2007.

[42] M. Müenchhof, M. Beck, and R. Isermann, "Fault-tolerant actuators and drives – Structures, fault detection principles and applications," *Annual Reviews in Control*, vol. 33, no. 2, pp. 136–148, 2009.

[43] N. Vaughan, P. Pomeroy, and D. Tilley, "The contribution of erosive wear to the performance degradation of sliding spool servovalves," *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, vol. 212, no. 6, pp. 437–451, 1998.

[44] R. W. Park, "Contamination control – A hydraulic OEM perspective," http://www.moog.com/literature/ICD/contaminationtechnicalarticle.pdf, 1997, online; accessed 28-June-2015.

[45] M. Münchhof, "Model-based fault detection for a hydraulic servo axis," Ph.D. dissertation, Technische Universität Darmstadt, 2006.

[46] M. Münchhof and M. Beck, "Model adjustment and multi-model based fault diagnosis for hydraulic servo axis," in *Proceedings of the 17th World Congress, the International Federation of Automatic Control, Seoul, South Korea*, 2008.

[47] L. An and N. Sepehri, "Leakage fault detection in hydraulic actuators subject to unknown external loading," *International Journal of Fluid Power*, vol. 9, no. 2, pp. 15–25, 2008.

[48] L. An and N. Sepehri, "Hydraulic actuator leakage quantification scheme using extended Kalman filter and sequential test method," in *American Control Conference, 2006.* IEEE, 2006, pp. 6–pp.

[49] M. Khoshzaban-Zavarehi, "On-line condition monitoring and fault diagnosis in hydraulic system components using parameter estimation and pattern classification," Ph.D. dissertation, University of British Columbia, 1997.

[50] Y. A. Chinniah, "Fault detection in the electrohydraulic actuator using extended Kalman filter," Ph.D. dissertation, University of Saskatchewan, Saskatoon, Canada, 2004.

[51] M. Sepasi and F. Sassani, "On-line fault diagnosis of hydraulic systems using unscented Kalman filter," *International Journal of Control, Automation and Systems*, vol. 8, no. 1, pp. 149–156, 2010.

[52] L. Chen, "Model-based fault diagnosis and fault-tolerant control for a nonlinear electro-hydraulic system," Ph.D. dissertation, Kaiserslautern, Techn. Univ., Diss., 2010, 2010.

[53] H. Khan, S. C. Abou, and N. Sepehri, "Nonlinear observer-based fault detection technique for electro-hydraulic servo-positioning systems," *Mechatronics*, vol. 15, no. 9, pp. 1037–1059, 2005.

[54] H.-Z. Tan and N. Sepehri, "Parametric fault diagnosis for electrohydraulic cylinder drive units," *IEEE Transactions on Industrial Electronics*, vol. 49, no. 1, pp. 96–106, 2002.

[55] T. Le, J. Watton, and D. Pham, "Fault classification of fluid power systems using a dynamics feature extraction technique and neural networks," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 212, no. 2, pp. 87–97, 1998.

[56] S. Mondal, G. Chakraborty, and K. Bhattacharyya, "Robust unknown input observer for nonlinear systems and its application to fault detection and isolation," *Journal of Dynamic Systems, Measurement, and Control*, vol. 130, no. 4, p. 044503, 2008.

[57] P. Garimella and B. Yao, "Fault detection of an electro-hydraulic cylinder using adaptive robust observers," in *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, 2004, pp. 119–128.

[58] P. Garimella and B. Yao, "Nonlinear adaptive robust observer design for a class of nonlinear systems," in *Proceedings of the American Control Conference*, vol. 5. IEEE, 2003, pp. 4391 – 4396.

[59] P. Garimella and B. Yao, "Model based fault detection of an electro-hydraulic cylinder," in *Proceedings of the American Control Conference*. IEEE, 2005, pp. 484 – 489.

[60] D. Yu, "Fault diagnosis for a hydraulic drive system using a parameter-estimation method," *Control Engineering Practice*, vol. 5, no. 9, pp. 1283 – 1291, 1997.

[61] D. Yu and D. Shields, "A bilinear fault detection observer," *Automatica*, vol. 32, no. 11, pp. 1597 – 1602, 1996.

[62] X. Wang and V. L. Syrmos, "Fault detection, identification and estimation in the electro-hydraulic actuator system using EKF-based multiple-model estimation," in *Proceedings of the 16th IEEE Mediterranean Conference on Control and Automation*, 2008, pp. 1693 – 1698.

[63] S. A. Gadsden, K. McCullough, and S. R. Habibi, "Fault detection and diagnosis of an electrohydrostatic actuator using a novel interacting multiple model approach," in *Proceedings of the American Control Conference*. IEEE, 2011, pp. 1396 – 1401.

[64] Z. Shi, F. Gu, B. Lennox, and A. Ball, "The development of an adaptive threshold for model-based fault detection of a nonlinear electro-hydraulic system," *Control Engineering Practice*, vol. 13, no. 11, pp. 1357 – 1367, 2005.

[65] K. Patan, M. Witczak, and J. Korbicz, "Towards robustness in neural network based fault diagnosis," *International Journal of Applied Mathematics and Computer Science*, vol. 18, no. 4, pp. 443 – 454, 2008.

[66] D. E. Whitney, "Resolved motion rate control of manipulators and human prostheses." *IEEE Transactions on Man-Machine Systems*, 1969.

[67] A. Liegeois, "Automatic supervisory control of the configuration and behavior of multibody mechanisms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 12, pp. 868 – 871, 1977.

[68] T. F. Chan and R. V. Dubey, "A weighted least-norm solution based scheme for avoiding joint limits for redundant joint manipulators," *IEEE Transactions on Robotics and Automation*, vol. 11, no. 2, pp. 286 – 292, 1995.

[69] A. R. Hirakawa and A. Kawamura, "Trajectory planning of redundant manipulators for minimum energy consumption without matrix inversion," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, 1997, pp. 2415 – 2420.

[70] A. Deo and I. Walker, "Minimum effort inverse kinematics for redundant manipulators," *IEEE Transactions on Robotics and Automation*, vol. 13, no. 5, pp. 767 – 775, 1997.

[71] M. Vukobratovic and M. Kircanski, "A dynamic approach to nominal trajectory synthesis for redundant manipulators," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-14, no. 4, pp. 580 – 586, July 1984.

[72] S. S. Rao, *Engineering optimization: Theory and practice.* John Wiley & Sons, 2009.

[73] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," Technical University of Denmark, Tech. Rep., 2004.

[74] "Ponsse Buffalo forwarder," http://www.ponsse.com/fi/tuotteet/kuormatraktorit/buffalo, 2017, [Online; accessed 28-February-2017].

[75] B. Siciliano, "Kinematic control of redundant robot manipulators: A tutorial," *Journal of Intelligent and Robotic Systems*, vol. 3, no. 3, pp. 201–212, 1990.

[76] K. Suh and J. M. Hollerbach, "Local versus global torque optimization of redundant manipulators," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 4, 1987, pp. 619–624.

[77] J. M. Hollerbach and K. Suh, "Redundancy resolution of manipulators through torque optimization," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 308–316, 1987.

[78] L. Beiner, "Minimum-force redundancy control of hydraulic cranes," *Mechatronics*, vol. 7, no. 6, pp. 537–547, 1997.

[79] Y. Zhang, S. S. Ge, and T. H. Lee, "A unified quadratic-programming-based dynamical system approach to joint torque optimization of physically constrained redundant manipulators," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 5, pp. 2126–2132, 2004.

[80] Y. Zhang and S. Ma, "Minimum-energy redundancy resolution of robot manipulators unified by quadratic programming and its online solution," in *Proceedings of the International Conference on Mechatronics and Automation (ICMA).* IEEE, 2007, pp. 3232–3237.

[81] Y. Zhang, D. Guo, and S. Ma, "Different-level simultaneous minimization of joint-velocity and joint-torque for redundant robot manipulators," *Journal of Intelligent & Robotic Systems*, vol. 72, no. 3-4, pp. 301–323, 2013.

[82] D. Ortiz Morales, S. Westerberg, P. X. La Hera, U. Mettin, L. Freidovich, and A. S. Shiriaev, "Increasing the level of automation in the forestry logging process with crane trajectory planning and control," *Journal of Field Robotics*, vol. 31, no. 3, pp. 343–363, 2014.

[83] U. Mettin, S. Westerberg, A. S. Shiriaev, and P. X. La Hera, "Analysis of human-operated motions and trajectory replanning for kinematically redundant manipulators," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2009, pp. 795–800.

[84] T. Balkan, "A dynamic programming approach to optimal control of robotic manipulators," *Mechanics Research Communications*, vol. 25, no. 2, pp. 225–230, 1998.

[85] K. G. Shin and N. D. McKay, "A dynamic programming approach to trajectory planning of robotic manipulators," *IEEE Transactions on Automatic Control*, vol. 31, no. 6, pp. 491–500, 1986.

[86] G. Field and Y. Stepanenko, "Iterative dynamic programming: An approach to minimum energy trajectory planning for robotic manipulators," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, 1996, pp. 2755–2760.

[87] A. R. Enes and W. J. Book, "Optimizing point to point motion of net velocity constrained manipulators," in *Proceedings of the 49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 6415–6420.

[88] P. Hołobut, "Time-optimal control of hydraulic manipulators with path constraints," *Journal of Theoretical and Applied Mechanics*, vol. 43, no. 3, pp. 523–538, 2005.

[89] U. Mettin, P. X. La Hera, D. Ortiz Morales, A. S. Shiriaev, L. B. Freidovich, and S. Westerberg, "Trajectory planning and time-independent motion control for a kinematically redundant hydraulic manipulator," in *Proceedings of the IEEE International Conference on Advanced Robotics (ICAR)*, 2009, pp. 1–6.

[90] B. Löfgren and J. Wikander, "Kinematic control of redundant knuckle booms," *International Journal of Forest Engineering*, vol. 20, no. 1, pp. 22–30, 2009.

[91] U. Mettin, "Principles for planning and analyzing motions of underactuated mechanical systems and redundant manipulators," Ph.D. dissertation, Umeå universitet, Institutionen för tillämpad fysik och elektronik, 2009.

[92] S. Ma and M. Watanabe, "Minimum time path-tracking control of redundant manipulators," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 1, 2000, pp. 27–32.

[93] K. G. Shin and N. D. McKay, "Minimum-time control of robotic manipulators with geometric path constraints," *IEEE Transactions on Automatic Control*, vol. 30, no. 6, pp. 531–541, 1985.

[94] J. E. Bobrow, S. Dubowsky, and J. Gibson, "Time-optimal control of robotic manipulators along specified paths," *The International Journal of Robotics Research*, vol. 4, no. 3, pp. 3–17, 1985.

[95] F. Pfeiffer and R. Johanni, "A concept for manipulator trajectory planning," *IEEE Journal of Robotics and Automation*, vol. 3, no. 2, pp. 115–123, 1987.

[96] D. Costantinescu and E. Croft, "Smooth and time-optimal trajectory planning for industrial manipulators along specified paths," *Journal of Robotic Systems*, vol. 17, no. 5, pp. 233–249, 2000.

[97] A. Gasparetto and V. Zanotto, "A technique for time-jerk optimal planning of robot trajectories," *Robotics and Computer-Integrated Manufacturing*, vol. 24, no. 3, pp. 415–426, 2008.

[98] M. H. Choi, "Redundancy resolution by minimization of joint disturbance torque for independent joint controlled manipulators," in *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 1999, pp. 392–397.

[99] H. Rodrigues, M. Monteiro, and D. Torres, "Optimal control and numerical software: An overview," *arXiv preprint arXiv:1401.7279*, 2014.

[100] M. Diehl, "Numerical optimal control - DRAFT," Optimization in Engineering Center (OPTEC) and ESAT-SCD, K.U. Leuven, Tech. Rep., 2011.

[101] J. L. Speyer and D. H. Jacobson, *Primer on optimal control theory.* SIAM, 2010, vol. 20.

[102] E. Todorov, "Optimal control theory," *Bayesian brain: Probabilistic approaches to neural coding*, pp. 269–298, 2006.

[103] Y. Nakamura and H. Hanafusa, "Optimal redundancy control of robot manipulators," *The International Journal of Robotics Research*, vol. 6, no. 1, pp. 32–42, 1987.

[104] D. P. Martin, J. Baillieul, and J. M. Hollerbach, "Resolution of kinematic redundancy using optimization techniques," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 4, pp. 529–533, 1989.

[105] S.-W. Kim, K.-B. Park, and J.-J. Lee, "Redundancy resolution of robot manipulators using optimal kinematic control," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 1994, pp. 683–688.

[106] R. Callies and P. Rentrop, "Optimal control of rigid-link manipulators by indirect methods," *GAMM-Mitteilungen*, vol. 31, no. 1, pp. 27–58, 2008.

[107] X. Wang, "Solving optimal control problems with Matlab: Indirect methods," ISE Dept., NCSU, Tech. Rep., 2009.

[108] J. T. Betts, *Practical methods for optimal control and estimation using nonlinear programming.* SIAM, 2010, vol. 19.

[109] V. M. Becerra, "Solving optimal control problems with state constraints using nonlinear programming and simulation tools," *IEEE Transactions on Education*, pp. 377–384, 2004.

[110] "Snopt sparse sequential quadratic programming algorithm," http://www.sbsi-sol-optimize.com/asp/sol_products_snopt_desc.htm, 2015, [Online; accessed 20-June-2015].

[111] "Knitro nonlinear optimisation solver," https://www.artelys.com/en/optimization-tools/knitro, 2015, [Online; accessed 19-March-2017].

[112] "Ipopt (Interior Point OPTimizer) nonlinear optimisation solver," https://projects.coin-or.org/Ipopt, 2015, [Online; accessed 20-June-2015].

[113] "Tomlab optimisation software," http://tomopt.com, 2015, [Online; accessed 20-June-2015].

[114] "GPOPS-II optimal control software," http://www.gpops2.com, 2015, [Online; accessed 20-June-2015].

[115] "ACADO Toolkit optimisation software," http://acado.github.io, 2015, [Online; accessed 19-March-2017].

[116] O. Sundstrom and L. Guzzella, "A generic dynamic programming Matlab function," in *Proc. of the IEEE Control Applications (CCA) & Intelligent Control (ISIC)*, St. Petersburg, Russia, 2009, pp. 1625–1630.

[117] T. J. Böhme and B. Frank, "Dynamic programming," in *Hybrid Systems, Optimal Control and Hybrid Vehicles*. Springer, 2017, pp. 199 – 214.

[118] J. Vihonen, J. Honkakorpi, J. Mattila, and A. Visa, "Geometry-aided MEMS motion state estimation for multi-body manipulators," in *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2013, pp. 341 – 347.

[119] J. Honkakorpi, "MEMS-based motion state estimation and control of hydraulic manipulators," Ph.D. dissertation, Tampere University of Technology, 2014.

Publications

Publication I

Jarmo Nurmi, Jouni Mattila "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 1: Global sensitivity analysis," *International Journal of Fluid Power*, vol. 12, no. 3, pp. 41 – 51, November 2011.

This is an Accepted Manuscript of an article published by Taylor & Francis in International Journal of Fluid Power on 11/2011, available online: http://www.tandfonline.com/10.1080/14399776.2011.10781036

DETECTION AND ISOLATION OF LEAKAGE AND VALVE FAULTS IN HYDRAULIC SYSTEMS IN VARYING LOADING CONDITIONS, PART 1: GLOBAL SENSITIVITY ANALYSIS

Jarmo Nurmi and Jouni Mattila

Tampere University of Technology, Department of Intelligent Hydraulics and Automation, P.O. Box 589, 33101 Tampere, Finland

E-mail: jarmo.nurmi@tut.fi

Abstract

Model-based condition monitoring methods are widely used in condition monitoring. They usually rely on ad hoc approaches to verify the system model and then best practices are reported to detect the given set of faults. This first part of a two-piece paper introduces a generic Global Sensitivity Analysis-based approach that can be applied systematically to verify the model parameter sensitivities used for the model-based fault detection. The case study is a generic servo valve-controlled hydraulic cylinder with unknown loading condition which is then systematically analyzed with Global Sensitivity Analysis. The method shows valuable insight into systematic model verification and resulting fault detection in terms of showing the dominant sensitivity of the nominal flow rate and nominal pressure difference, and the exact sensitivities of 0-1 dm^3/min external and internal leakages on cylinder chamber pressures and velocity. In the second paper, an Unscented Kalman Filter-based Fault Detection and Isolation scheme for leakage and valve faults of a generic servo valve-controlled hydraulic cylinder is devised and fault patterns are presented.

Keywords: global sensitivity analysis, fault detection, model verification, Sobol' indices

1 Introduction

In a fault detection process, system condition is constantly observed and decisions are made whether the system has faults. Once a fault is detected, a fault isolation process takes over and localizes the cause of the fault.

To avoid false alarms (or false positives), in model-based condition monitoring it is important to verify the model to be as accurate as possible. A Global Sensitivity Analysis (GSA, Saltelli et al. 2008) helps in the verification, since it reveals the most sensitive parameters of the system in a systematic way. By focusing efforts on improving the sensitive parameters, a more robust model is reached. The GSA can also reveal the sensitivity of faults on system outputs, which is useful for fault detection purposes.

Model sensitivity can be analyzed locally. For instance, a Local Sensitivity Analysis (LSA) to a nonlinear variable displacement axial piston pump model was applied to study parameter sensitivities and to reduce model order (Kim et al., 1987), and to a linear water quality model (Pastres et al., 1997). In LSA, parameters are deviated individually from their nominal values, which can be performed analytically with Eq. (1) if the model output is differentiable and otherwise numerically with Eq. (2):

$$S_i = \frac{\partial Y}{\partial X_i}, \qquad i = \{1,2,3,\dots,k\} \tag{1}$$

$$S_i = \frac{Y - f(X_1, X_2, X_3, \dots, X_i + \Delta X_i, \dots, X_k)}{\Delta X_i} \tag{2}$$

where S_i is the sensitivity of output $Y = f(X_1, X_2, X_3, \dots, X_k)$ to a change in parameter X_i.

Thus LSA sensitivities are valid in close proximity of nominal parameters. Therefore, GSA is more applicable to nonlinear models since GSA sensitivities are valid in a wider parameter space. Previously, GSA has been used in studying, for example the sensitive forces in a pipe bend and parameters in a dam-break experiment (Hall. et al., 2009), and the parameters in water hammer model (Kaliatka et al., 2009). The GSA method of this paper is the variance-based Sobol' indices because it is simpler to implement than for instance the Fourier Amplitude Sensitivity Test (FAST).

Current studies on model-based condition monitoring rely on an ad hoc approach to find the best ways to detect faults and to verify the model parameters. Our proposal in this paper is the systematic utilization of GSA to verify the system model and to find the best practices to detect faults.

This paper is organized as follows. In Section 2, the mechanism of a generic valve-controlled hydraulic cylinder that drives a manipulator joint is presented. Then the corresponding test bed is introduced and modelled. In Section 3, the GSA algorithm and its implementation with Monte Carlo methods are described. In Section 4, the GSA is applied to the test bed and the results are discussed.

2 Modelling and Test Bed

The objective of the fault detection and isolation scheme that is devised in part 2 on the basis of part 1 is that it is applicable to a generic valve-controlled cylinder that drives any of the n-DOF manipulator joints, see Fig. 1:.

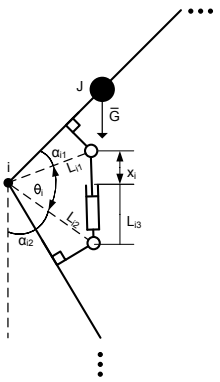


Fig. 1: *A manipulator joint driven by a hydraulic cylinder.*

The piston position of the ith cylinder is given by the law of cosines:

$$x_{\mathrm{i}}(\theta_{\mathrm{i}}) = \sqrt{L_{\mathrm{i1}}^2 + L_{\mathrm{i2}}^2 - 2L_{\mathrm{i1}}L_{\mathrm{i2}}\cos\theta_{\mathrm{i}}} - L_{i3} \tag{3}$$

Piston velocity of the ith cylinder can be differentiated from Eq. (3):

$$\dot{x}_{\mathrm{i}}(\theta_{\mathrm{i}}) = r_{\mathrm{i}}(\theta_{\mathrm{i}})\dot{\theta}_{\mathrm{i}} \tag{4}$$

where $\dot{\theta}_{\mathrm{i}}$ is the angular velocity of the joint and the torque arm of the ith cylinder is given by:

$$r_{\mathrm{i}}(\theta_{\mathrm{i}}) = \frac{L_{\mathrm{i1}}L_{\mathrm{i2}}\sin\theta_{\mathrm{i}}}{\sqrt{L_{\mathrm{i1}}^2 + L_{\mathrm{i2}}^2 - 2L_{\mathrm{i1}}L_{\mathrm{i2}}\cos\theta_{\mathrm{i}}}} \tag{5}$$

Consider an open chain manipulator system that consists of n cylinders. The piston velocities of all cylinders can be presented compactly with matrix notation:

$$\dot{\boldsymbol{x}} = \mathbf{R}(\boldsymbol{\theta})\dot{\boldsymbol{\theta}} = \begin{bmatrix} \mathrm{r}_1(\theta_1) & 0 & \cdots & 0 \\ 0 & \mathrm{r}_2(\theta_2) & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{r}_{\mathrm{n}}(\theta_{\mathrm{n}}) \end{bmatrix} \dot{\boldsymbol{\theta}} \tag{6}$$

The torques acting on the joints expressed with linear actuator coordinates are (Beiner and Mattila, 1999):

$$\begin{aligned} \boldsymbol{\tau}_{\mathrm{cyl}} &= \mathbf{R}(\boldsymbol{\theta})\boldsymbol{F} \\ &= \mathbf{J}(\boldsymbol{\theta})\mathbf{R}(\boldsymbol{\theta})^{-1}\ddot{\boldsymbol{x}} - \mathbf{J}(\boldsymbol{\theta})\mathbf{R}(\boldsymbol{\theta})^{-1}\dot{\mathbf{R}}(\boldsymbol{\theta})\mathbf{R}(\boldsymbol{\theta})^{-1}\dot{\boldsymbol{x}} + \boldsymbol{V}(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}) \\ &\quad + \boldsymbol{G}(\boldsymbol{\theta}) \end{aligned} \tag{7}$$

where $\mathbf{R}(\boldsymbol{\theta})\boldsymbol{F}$ consists of cylinder actuator torques, $\boldsymbol{V}(\boldsymbol{\theta},\dot{\boldsymbol{\theta}})$ consists of torques caused by the Coriolis effect and centrifugal force, and $\boldsymbol{G}(\boldsymbol{\theta})$ is the vector of gravitational torques.

2.1 Case Study –Test Bed

The GSA is applied to a hydraulic boom called Single Axis Mock-up (SAM), shown in Fig. 2:.

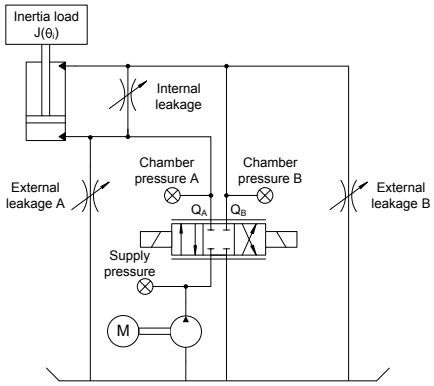


Fig. 2: *The hydraulic diagram of the SAM.*

The SAM has a 4/3-directional valve that controls the joint cylinder. Three restrictor valves are used to emulate external leakages ('External leakage A' and 'External leakage B') and internal leakage ('Internal leakage'). The external leakage emulates fluid leakage to the environment due to a broken hose, pipe or a failed coupling, while the internal leakage arrangement emulates cylinder seal failure. The system components are listed to Appendix 1, Table 9:. The SAM, with a 4.5 Hz maximum hydraulic natural frequency (Fig. 4:), is illustrated in Fig. 3:.

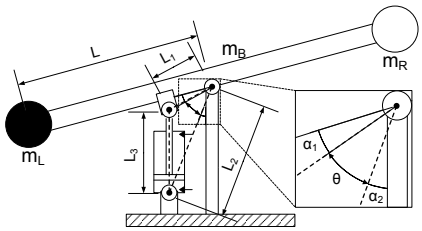


Fig. 3: *An illustration of the boom.*

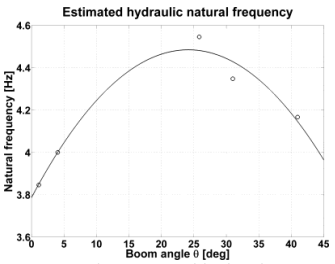


Fig. 4: *The estimated hydraulic natural frequency of SAM.*

2.2 Case Study –Test Bed Model

The mathematical model is divided into hydraulic system equations, motion equations of the boom, and then the entire model is presented in a continuous-time and in a discretized state space form.

2.2.1 Hydraulic System Equations

Equations (8) and (9) describing the change in chamber pressures are as follows (Watton, 1989):

$$\dot{p}_A = \frac{B_{effA}}{V_{0A} + A_A x}(Q_A(\mathrm{p_A}, \mathrm{x_s}) - A_A \dot{x}) \tag{8}$$

$$\dot{p}_B = \frac{B_{effB}}{V_{0B} + A_B(x_{max} - x)}(Q_B(\mathrm{p_B}, \mathrm{x_s}) + A_B \dot{x}) \tag{9}$$

where B_{effX} is the effective bulk modulus in chamber X (for X = A, B), $V_{0\mathrm{X}}$ is the volume in chamber X, A_{X} is the area in chamber X, Q_{X} is the flow sum to and from chamber X, $x_{\max}$ is the cylinder stroke, x is the piston position and $\dot{x}$ denotes velocity.

The algebraic equations for flows $Q_{\mathrm{A}}(\mathrm{p_A}, \mathrm{x_s})$ and $Q_{\mathrm{B}}(\mathrm{p_B}, \mathrm{x_s})$ with the flow into the cylinder being positive can be written as follows:

$$Q_A(\dots) = \begin{cases} K_{vPA}(x_s + offset)\sqrt{|p_S - p_A|}sgn(p_S - p_A) + Q_{leakA}, & x_s + offset > 0 \\ -K_{vAT}(-x_s + offset)\sqrt{|p_A - p_T|}sgn(p_A - p_T) + Q_{leakA}, & x_s + offset < 0 \\ 0, & x_s = -offset \end{cases} \tag{10}$$

$$Q_B(\dots) = \begin{cases} -K_{vBT}(x_s + \mathrm{offset})\sqrt{|p_B - p_T|}\mathrm{sgn}(p_B - p_T) + Q_{leakB}, & x_s + \mathrm{offset} > 0 \\ K_{vPB}(-x_s + \mathrm{offset})\sqrt{|p_S - p_B|}\mathrm{sgn}(p_S - p_B) + Q_{leakB}, & x_s + \mathrm{offset} < 0 \\ 0, & x_s = -offset \end{cases} \tag{11}$$

where $\mathrm{K_{vX}}$ is flow coefficient in notch X, for X = PA, AT, BT and PB, x_{s} is the spool position, offset denotes the deviation of the valve spool from its correct position, p_{S} is the supply pressure, p_{A} is the pressure A, p_{T} is the tank pressure and p_{B} is the pressure B. The flow coefficients are defined as follows:

$$K_{\mathrm{vX}} = \frac{Q_{N,X}}{\sqrt{\Delta \mathrm{p_{N,X}}}} \tag{12}$$

where $Q_{\mathrm{N,X}}$ is the nominal flow rate and $\Delta \mathrm{p_{N,X}}$ is the nominal pressure difference in notch X.

The terms Q_{leakA} and Q_{leakB} are laminar leakage flows, present when the spool position is between -1 % and 1 % of its maximum:

$$\begin{cases} Q_{leakA} = K_{vPA,leak}(p_S - p_A) - K_{vAT,leak}(p_A - p_T), & |\mathrm{x_s}| < 0.01 \\ 0, & otherwise \end{cases} \tag{13}$$

$$\begin{cases} Q_{leakB} = K_{vPB,leak}(p_S - p_B) - K_{vBT,leak}(p_B - p_T) & |\mathrm{x_s}| < 0.01 \\ 0, & otherwise \end{cases} \tag{14}$$

where $K_{vX,leak}$ are the leakage flow coefficients.

The spool x_s dynamics are modelled with the 2nd order differential equation:

$$\ddot{x}_s = K\omega_n^2 u - 2\omega_n d_r \dot{x}_s - \omega_n^2 x_s \tag{15}$$

2.2.2 Motion Equations of the Boom

The piston position $x(\theta)$ is calculated according to Eq. (3) and the velocity $\dot{x}(\theta)$ according to Eq. (4).

We calculate the angular acceleration of the boom $\ddot{\theta}$ by dividing the sum of torques acting on the boom with total moment of inertia as follows:

$$\ddot{\theta} = \frac{\sum \overline{\tau}}{J_{\mathrm{tot}}} = \frac{\tau_{\mathrm{cyl}} + \tau_{\mathrm{mR}} - \tau_{\mathrm{mL}} - \tau_{\mathrm{B}}}{\frac{1}{12} m_{\mathrm{B}} {L_{\mathrm{B}}}^2 + m_{\mathrm{L}} L^2 + m_{\mathrm{R}} L^2} \tag{16}$$

where τ_{cyl} is the torque generated by the cylinder, τ_{mR} and τ_{mL} are the torques caused by the load masses on the right and left, respectively, and τ_{B} is the torque produced by the boom, since the boom is not jointed to the base from its center of gravity. The total moment of inertia J_{tot} consists of the load masses m_{L} and m_{R} at a distance L from the center of rotation, and of the boom's moment of inertia with mass m_{B} and length L_{B}.

The friction force F_{μ} is as follows (Canudas de Wit et al., 1995):

$$F_{\mu}(\dot{x}, \mathrm{z}) = \sigma_0 z + \sigma_1 \underbrace{[\frac{\sigma_0 |\dot{x}|}{[\mathrm{F_C} + (\mathrm{F_S} - \mathrm{F_C})\mathrm{e}^{-(\dot{x}/v_s)^2}}\mathrm{z}]}_{\dot{z}} + b\dot{x} \tag{17}$$

where z is the bending of the cylinder seal, σ_0 is the stiffness of the seal, σ_1 is the damping coefficient and b is the viscous friction coefficient. This friction model includes the stick-slip phenomenon. For more information on the dynamics of state variable z refer to the original publication (Canudas de Wit et al., 1995). The relation between pressure levels and friction force was neglected.

2.2.3 State Space Representation of the Model

The entire model can be presented compactly in state space form. The states of the system are:

$$\boldsymbol{x} = [p_{\mathrm{A}}, p_{\mathrm{B}}, x_{\mathrm{s}}, \dot{x}_{\mathrm{s}}, x, \dot{x}, z]^{\mathrm{T}} = [x_1, x_2, x_3, x_4, x_5, x_6, x_7]^{\mathrm{T}} \tag{18}$$

The continuous-time state space representation is then:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix} = \begin{bmatrix} \frac{B_{\mathrm{effA}}}{V_{0\mathrm{A}} + A_{\mathrm{A}} x_5}(Q_{\mathrm{A}}(\mathrm{x_1}, \mathrm{x_3}) - A_{\mathrm{A}} x_6) \\ \frac{B_{\mathrm{effB}}}{V_{0\mathrm{B}} + A_{\mathrm{B}}(x_{\max} - x_5)}(Q_{\mathrm{B}}(\mathrm{x_2}, \mathrm{x_3}) + A_{\mathrm{B}} x_6) \\ x_4 \\ K\omega_{\mathrm{n}} u - 2\omega_{\mathrm{n}} d_{\mathrm{r}} x_4 - \omega_{\mathrm{n}}^2 x_3 \\ x_6 \\ m_{\mathrm{reduced}}^{-1}(x_1 A_{\mathrm{A}} - x_2 A_{\mathrm{B}} - F_{\mu}(\mathrm{x_6}, \mathrm{x_7}) - F_{\mathrm{ext}}) \\ x_6 - \sigma_0 |x_6| [F_{\mathrm{c}} + (F_{\mathrm{S}} - F_{\mathrm{c}}) e^{-(x_6/v_{\mathrm{s}})^2}]^{-1} x_7 \end{bmatrix} \tag{19}$$

where m_{reduced} is the reduced mass on the cylinder and F_{ext} is the external force which can be written as:

$$m_{\text{reduced}}(\theta) = \frac{J_{\text{tot}}}{r^2(\theta)} \tag{20}$$

$$F_{\text{ext}} = \frac{-\tau_{\text{mR}} + \tau_{\text{mL}} + \tau_{\text{B}}}{r} \tag{21}$$

The continuous-time state space representation can be transformed to discrete-time with sampling time T with Euler's forward method:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \\ x_5(k+1) \\ x_6(k+1) \\ x_7(k+1) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \\ x_7(k) \end{bmatrix} +$$

$$T \begin{bmatrix} \frac{B_{\text{effA}}}{A_{\text{A}} x_5(k) + V_{0\text{A}}} (Q_{\text{A}}(\mathrm{x}_1(k), \mathrm{x}_3(k)) - A_{\text{A}} x_6(k)) \\ \frac{B_{\text{effB}}}{A_{\text{A}}(x_{\max} - x_5(k)) + V_{0\text{B}}} (Q_{\text{B}}(\mathrm{x}_2(k), \mathrm{x}_3(k)) + A_{\text{B}} x_6(k)) \\ x_4(k) \\ K\omega_{\text{n}} u(k) - 2\omega_{\text{n}} d_r x_4(k) - \omega_{\text{n}}^2 x_3(k) \\ x_6(k) \\ m_{\text{reduced}}^{-1}(x_1(k) A_{\text{A}} - x_2(k) A_{\text{B}} - F_{\mu}(\mathrm{x}_6(k), \mathrm{x}_7(k)) - F_{\text{ext}}) \\ x_6(k) - \sigma_0 |x_6(k)| \left[F_{\text{c}} + (F_{\text{S}} - F_{\text{c}}) e^{-\left(\frac{x_6(k)}{v_{\text{s}}}\right)^2} \right]^{-1} x_7(k) \end{bmatrix} \tag{22}$$

3 Global Sensitivity Analysis

A Global Sensitivity Analysis (GSA) method called Sobol' indices, its computation procedure and its usefulness for condition monitoring and model verification are introduced in this section.

3.1 Sobol' Indices Method

The premises for the variance-based Sobol' indices method are as follows (Saltelli et al., 2008, pp. 160-163). Consider the model to be a square-integrable function $\boldsymbol{Y} = \mathrm{f}(\boldsymbol{X})$ which can be divided into summands of increasing dimensionality:

$$\begin{aligned} &\mathrm{f}(X_1, X_2, X_3, \dots X_{\mathrm{k}}) \\ &= \mathrm{f}_0 + \sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{f}_{\mathrm{i}}(X_{\mathrm{i}}) \\ &+ \sum_{\mathrm{i}=1}^{\mathrm{k}} \sum_{\mathrm{j}=\mathrm{i}+1}^{\mathrm{k}} \mathrm{f}_{\mathrm{ij}}(X_{\mathrm{i}}, X_{\mathrm{j}}) + \cdots \\ &+ \mathrm{f}_{1\dots\mathrm{k}}(X_1, X_2, X_3, \dots X_{\mathrm{k}}) \end{aligned} \tag{23}$$

where k is the number of parameters and X_{k} is the random parameter k.

Equation (23) has a total of 2^{k} terms and infinite solutions. Sobol' proposed one solution, which decomposes the function $\mathrm{f}(\boldsymbol{X})$ into conditional expectations. The first three terms can be written as:

$$\mathrm{f}_0 = \mathrm{E}(\boldsymbol{Y}) \tag{24}$$

$$\mathrm{f}_{\mathrm{i}} = \mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}}) - \mathrm{E}(\boldsymbol{Y}) \tag{25}$$

$$\mathrm{f}_{\mathrm{ij}} = \mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}}, X_{\mathrm{j}}) - \mathrm{f}_{\mathrm{i}} - \mathrm{f}_{\mathrm{j}} - \mathrm{E}(\boldsymbol{Y}) \tag{26}$$

where $\mathrm{E}(\boldsymbol{Y})$ denotes the expectation of model output $\boldsymbol{Y}$, $\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})$ is the conditional expectation of output $\boldsymbol{Y}$ given that input X_{i} is fixed to a certain value.

The variances of Eq. (24)-(26) have the following properties:

$$\mathrm{V}_0 = \mathrm{V}(\mathrm{f}_0) = 0 \tag{27}$$

$$\mathrm{V}_{\mathrm{i}} = \mathrm{V}(\mathrm{f}_{\mathrm{i}}) = \mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})] \tag{28}$$

$$\begin{aligned} &\mathrm{V}_{\mathrm{ij}} = \mathrm{V}(\mathrm{f}_{\mathrm{ij}}) \\ &= \mathrm{V}\big[\mathrm{E}\big(\boldsymbol{Y}\big|X_{\mathrm{i}}, X_{\mathrm{j}}\big)\big] - \mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})] - \mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{j}})] \end{aligned} \tag{29}$$

The conditional variance in Eq. (28) is used to calculate first order sensitivity indices, which is a measure on the main effect of parameter X_{i} on output $\boldsymbol{Y}$:

$$\mathrm{S}_{\mathrm{i}} = \frac{\mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})]}{\mathrm{V}(\boldsymbol{Y})} \tag{30}$$

where $\mathrm{V}(\boldsymbol{Y})$ is the unconditional variance of output $\boldsymbol{Y}$.

The interpretation for the term $\mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})]$ is that first the conditional expectation $\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})$ is calculated by fixing the input X_{i} to a certain value X_{i}^* and allowing other inputs to vary. Thus a complete representation for the conditional expectation is $\mathrm{E}_{\sim \mathrm{X}_{\mathrm{i}}}(\boldsymbol{Y}|X_{\mathrm{i}} = X_{\mathrm{i}}^*)$. The $\sim X_{\mathrm{i}}$ operator means that the expectation is calculated over every input excluding X_{i}. For k inputs that is a set $\{X_1, X_2, \dots, X_{\mathrm{i}-1}, X_{\mathrm{i}+1}, \dots, X_{\mathrm{k}}\}$. Then the variance of the expectation is calculated over different values of X_{i}. Thus the complete representation for the nominator is $\mathrm{V}_{\mathrm{X}_{\mathrm{i}}}[\mathrm{E}_{\sim \mathrm{X}_{\mathrm{i}}}(Y|X_{\mathrm{i}} = X_{\mathrm{i}}^*)]$.

The total order effects, which include the first order effect but also the terms that come from interaction between parameters, is derived from the law of total variance:

$$\mathrm{V}(\boldsymbol{Y}) = \mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\sim \mathrm{i}})] + \mathrm{E}[\mathrm{V}(\boldsymbol{Y}|X_{\sim \mathrm{i}})] \tag{31}$$

where the first term is the main effect and the remaining term the residual.

Then the total order sensitivity indices can be calculated with:

$$\mathrm{S}_{\mathrm{T}_{\mathrm{i}}} = \frac{\mathrm{E}[\mathrm{V}(\boldsymbol{Y}|X_{\sim \mathrm{i}})]}{\mathrm{V}(\boldsymbol{Y})} = 1 - \frac{\mathrm{V}[E(\boldsymbol{Y}|X_{\sim \mathrm{i}})]}{\mathrm{V}(\boldsymbol{Y})} \tag{32}$$

where the latter equality is obtained from Eq. (31) by solving it for $\mathrm{E}[\mathrm{V}(\boldsymbol{Y}|X_{\sim \mathrm{i}})]$ and placing it to the former equality.

In Eq. (32), $\mathrm{E}[\mathrm{V}(\boldsymbol{Y}|X_{\sim \mathrm{i}})]$ is a term that contains the variance of $\boldsymbol{Y}$ that would be left if all inputs but X_{i} could be fixed (X_{i} would be allowed to vary). Thus, diving $\mathrm{E}[\mathrm{V}(\boldsymbol{Y}|X_{\sim \mathrm{i}})]$ by $\mathrm{V}(\boldsymbol{Y})$ gives the proportion of the variance that is caused by X_{i}. The term $\mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\sim \mathrm{i}})]$ contains the variance that would disappear from $\mathrm{V}(\boldsymbol{Y})$ if all inputs but X_{i} could be fixed.

3.2 Computing Sobol' Indices

The analytical computation of Sobol' indices of differential equation models is not possible. From Saltelli (2002) and Saltelli et al. (2008, pp. 164-167) a procedure for computing sensitivity indices is reached.

Consider two matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ which are of size N x k. N is the sample size (the number of simulations), and k is the amount of parameters that are randomly varied. In $\mathbf{X}_1$ and $\mathbf{X}_2$ each row m, for example $\boldsymbol{X}_1^{\mathrm{m}}$ with $\mathrm{m} = \{1,2,3,\dots,\mathrm{N}\}$, corresponds to one simulation with k random inputs in the columns. The inputs are varied using quasi-random numbers from the Sobol' sequence (Sobol' & Kucherenko, 2005), which produces more accurate sensitivity indices than pseudorandom numbers drawn randomly.

Simulating with input matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ N times, two output matrices $\mathbf{Y}_1 = \mathrm{f}(\mathbf{X}_1)$ and $\mathbf{Y}_2 = \mathrm{f}(\mathbf{X}_2)$ of size N x M are created, where M is the number of outputs. For $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ the estimated variances are:

$$\widehat{\mathrm{V}}(\mathbf{Y}_1) = \frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}^2(\boldsymbol{X}_1^{\mathrm{m}}) - \hat{\mathrm{f}}_1^2 \tag{33}$$

$$\widehat{\mathrm{V}}(\mathbf{Y}_2) = \frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}^2(\boldsymbol{X}_2^{\mathrm{m}}) - \hat{\mathrm{f}}_2^2 \tag{34}$$

where the squared expectation estimates are:

$$\hat{\mathrm{f}}_1^2 = \frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}(\boldsymbol{X}_1^{\mathrm{m}}) f(\boldsymbol{X}_2^{\mathrm{m}}) \tag{35}$$

$$\hat{\mathrm{f}}_2^2 = (\frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}(\boldsymbol{X}_2^{\mathrm{m}}))^2 \tag{36}$$

Equation (33) and Eq. (34) are used for computing first and total order indices, respectively.

We introduce input matrix $\mathbf{X}_{3\mathrm{i}}$ for calculating the nominators in Eq. (30) and (32). Matrix $\mathbf{X}_{3\mathrm{i}}$ has the same values as $\mathbf{X}_2$ except that the i^{th} column is taken from matrix $\mathbf{X}_1$. The first order sensitivity indices are:

$$\mathrm{S}_{\mathrm{i}} = \frac{\mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\mathrm{i}})]}{\mathrm{V}(\boldsymbol{Y})} = \frac{\frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}(\boldsymbol{X}_1^{\mathrm{m}})\,\mathrm{f}(\boldsymbol{X}_{3\mathrm{i}}^{\mathrm{m}}) - \hat{\mathrm{f}}_1^2}{\frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}^2(\boldsymbol{X}_1^{\mathrm{m}}) - \hat{\mathrm{f}}_1^2} \tag{37}$$

In the scalar product $\mathrm{f}(\boldsymbol{X}_1^{\mathrm{m}})\mathrm{f}(\boldsymbol{X}_{3\mathrm{i}}^{\mathrm{m}})$ the columns for X_{i} are the same. If X_{i} is influential, high and low values of outputs $\mathrm{f}(\boldsymbol{X}_1^{\mathrm{m}})$ and $\mathrm{f}(\boldsymbol{X}_{3\mathrm{i}}^{\mathrm{m}})$ are associated (a high value multiplied by a high value or a low value multiplied by a low value) and thus produce a higher value for the variance when the terms in the scalar product are added together. If X_i is a non-influential input, the high and low values of $\mathrm{f}(\boldsymbol{X}_1^{\mathrm{m}})$ and $\mathrm{f}(\boldsymbol{X}_{3\mathrm{i}}^{\mathrm{m}})$ are randomly associated, thus resulting in a lower value for the nominator.

The total order sensitivity indices are:

$$\begin{aligned}\mathrm{S}_{\mathrm{T}_{\mathrm{i}}} &= 1 - \frac{\mathrm{V}[\mathrm{E}(\boldsymbol{Y}|X_{\sim \mathrm{i}})]}{\mathrm{V}(\boldsymbol{Y})} \\ &= 1 - \frac{\frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}(\boldsymbol{X}_2^{\mathrm{m}})\,\mathrm{f}(\boldsymbol{X}_{3\mathrm{i}}^{\mathrm{m}}) - \hat{\mathrm{f}}_2^2}{\frac{1}{N}\sum_{\mathrm{m}=1}^{\mathrm{N}} \mathrm{f}^2(\boldsymbol{X}_2^{\mathrm{m}}) - \hat{\mathrm{f}}_2^2}\end{aligned} \tag{38}$$

The explanation for Eq. (38) is that as values for $X_{\sim \mathrm{i}}$ are the same and only the input X_{i} is randomly varied, if X_{i} is influential, the values in the product $\mathrm{f}(\boldsymbol{X}_2^{\mathrm{m}})\mathrm{f}(\boldsymbol{X}_{3\mathrm{i}}^{\mathrm{m}})$ will be randomly associated and produce a lower value in the latter term. But taking into consideration that this low value is subtracted from one, a high value will be the result, thus indicating that this input is meaningful, and vice versa.

At the expense of increased computational costs increasing sample size N results in better sensitivity index estimates, the described method requires N(2+k) model runs.

3.3 Interpreting Sensitivity Indices from a Condition Monitoring Perspective

Sensitivity indices are used to rank parameters according to their sensitivities for model verification purposes (Saltelli et al. 2008, pp. 166-167):

- Inputs with the lowest total order sensitivity indices (near zero) causing the least variance to the output can be fixed at a value between their examined bounds without compromising the accuracy of the model.
- Inputs with the highest first order sensitivity indices should be a priority in model verification, because their correct values will reduce the variance in output Y the most.

Model parameter interactions can be studied too. The interactions mean that the effect of parameter changes on the output is different if the parameters are changed together as opposed to individually changing them and summing their effects. The differences $S_{\mathrm{T}_{\mathrm{i}}} -$

S_i and $1 - \sum_{i=1}^{k} S_i$ are direct measures of the interactions. They are zero for perfectly additive models but nonzero for non-additive models.

GSA results are useful for fault detection purposes. The sensitivity indices of fault parameters indicate if the fault can be detected, and at which output.

4 Global Sensitivity Analysis of the Single Axis Mock-up

The sensitivity of the system for parameter changes and leakage faults is studied in this section, both in transients and in steady state. The main objective was to extract information from the sensitivities for fault detection purposes.

The valve control signal was a step signal to 25 % opening. Only the extending movement of the cylinder was examined because we wanted to limit the range of the study. Moreover, the asymmetry of the cylinder might affect the results in retraction. The sample size N in both analyses was 10000, with a fixed 1-millisecond simulation step size. The examined outputs were pressures A, B and velocity.

4.1 Sensitivity for Parameter Changes

The sensitivity of the SAM for seven varying parameters (k = 7) and their respective ranges are examined (Table 1:) using the model in Eq. (19).

Table 1: Single Axis Mock-up parameters deviated in the GSA.

Parameter	Explanation	Lower bound	Upper bound
offset	Spool deviation from actual position	-5 %	5 %
$Q_{N,PA}$	Nominal flow rate in notch PA	15 L/min	35 L/min
$\Delta p_{N,PA}$	Nominal pressure difference in notch PA	5 bar	40 bar
$Q_{N,BT}$	Nominal flow rate in notch BT	15 L/min	35 L/min
$\Delta p_{N,BT}$	Nominal pressure difference in notch BT	5 bar	40 bar
B_{eff}	Effective bulk modulus	300 MPa	1200 MPa
b	Viscous friction coefficient	2000 Ns/m	5000 Ns/m

The parameters are assumed to be uniformly distributed. The lower and upper bounds are chosen so that they are reasonable. For instance, the nominal flow rate of the valve is 24 L/min, thus a 15-35 L/min range is suitable. The nominal pressure differences are chosen so that most valve types fall within the range. Valve offset is a calibration error or a deviation caused by a valve fault. The constant parameters that were identified, measured or taken from manufacturer data are listed in Appendix 2.

The first order indices in steady state (Fig. 5:), including errors bounds calculated with a re-sampling method (bootstrapping), Archer et al. (1997), show how much variance in pressures and velocity is caused by individual parameters alone. The effective bulk modulus causes minor output variance. This is entirely intuitive because B_{eff} is a parameter that affects the natural frequency of the system and only has an effect on pressures or velocity in transients. This can be verified by setting either one of the pressure differential equations to zero to find the steady state pressures. Viscous friction coefficient b has insignificant magnitude, as indicated by the negligible sensitivity index b.

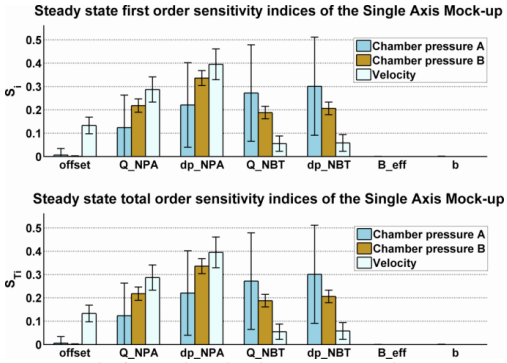


Fig. 5: *The first and total order sensitivity indices in steady state.*

Table 2: The ranking of parameters according to their first order indices.

Ranking	p_A	p_B	v
1	$\Delta p_{N,BT}$	$\Delta p_{N,PA}$	$\Delta p_{N,PA}$
2	$Q_{N,BT}$	$Q_{N,PA}$	$Q_{N,PA}$
3	$\Delta p_{N,PA}$	$\Delta p_{N,BT}$	*offset*
4	$Q_{N,PA}$	$Q_{N,BT}$	$\Delta p_{N,BT}$
5	*offset*	*offset*	$Q_{N,BT}$
6	B_{eff}	B_{eff}	B_{eff}
7	b	b	b

Table 2: ranks the parameters from Fig. 5:. It shows that the pressure variance is mostly captured by nominal pressure differences and flow rates. Notch BT parameters cause the most variance to pressure p_A, where as notch PA parameters are responsible for most of the variance to pressure p_B. The reason becomes clear from the steady state pressures:

$$p_{ssA} = \frac{F_{ext}A_A A_B K_{vBT}^2 + A_B^3 K_{vPA}^2 p_S + A_A^2 A_B K_{vBT}^2 p_T}{A_A^3 K_{vBT}^2 + A_B^3 K_{vPA}^2} \quad (39)$$

$$p_{ssB} = \frac{-F_{ext}K_{vPA}^2 A_B^2 + A_A A_B^2 K_{vPA}^2 p_S + A_A^3 K_{vBT}^2 p_T}{A_A^3 K_{vBT}^2 + A_B^3 K_{vPA}^2} \quad (40)$$

The steady state Eq. (39) can be derived by setting the pressure differential Eq. in (8) and (9) to zero and solving both for velocity. Then equating the resulting equations, replacing p_B with p_A calculated from the steady state motion equation of the piston, in Eq. (19), and finally solving for p_A and assuming that the friction force is included in the external force F_{ext} gives Eq. (39). Equation (40) is obtained likewise.

A difference in the steady state pressure equations is that in Eq. (39) F_{ext} is multiplied by flow coefficient K_{vBT}^2 and by K_{vPA}^2 in Eq. (40). Therefore, it is clear that parameters in notch BT affect pressure A more than pressure B. Similarly, parameters in notch PA cause more variance to pressure B. The nominal pressure differences are more influential than nominal flow rate because nominal pressure differences are varied along a wider range.

The sensitivity indices show that valve offset is influential on steady state velocity, which is obvious since the offset affects valve opening. The ranks for rest of the parameters affecting velocity are fairly intuitive.

For a measure of interactions between parameters, we sum up the first order indices of each parameter for each output. The results are presented in Table 3:. The interactions among parameters are negligible for each output, which means that the total order indices (Fig. 5:) do not differ remarkably from the first order indices. The unexplained part is five to seven percent in each output, which could be caused by estimation errors in the calculations. Increasing sample size could possibly reduce this.

Table 3: Parameter interactions in steady state.

Output	Interaction measure: $1 - \sum_{i=1}^{k} S_i$
p_A	0.0763
p_B	0.0507
v	0.0713

For computation of sensitivity indices in transients, an area plot is illustrative and the amount of interactions between parameters is visible. A general rule of thumb for reading the area plot is that the bigger the area of the parameter, the larger its effect is. Figure 6 presents interpolated first and total order indices for pressure A computed at time instants 0.05, 0.10, 0.20, 0.30, 0.5, 0.75, 1.0, 2.0, 3.0 and 4.0 seconds.

The first order indices in Fig. 6: show a sensitivity drop at 0.10 seconds to about 0.30 seconds. Particularly, the indices of valve nominal parameters in notch PA and the nominal flow rate of notch BT drop significantly. The figure shows that these are more significant parameters, especially at the beginning of the motion. After about a second all indices reach their steady state level.

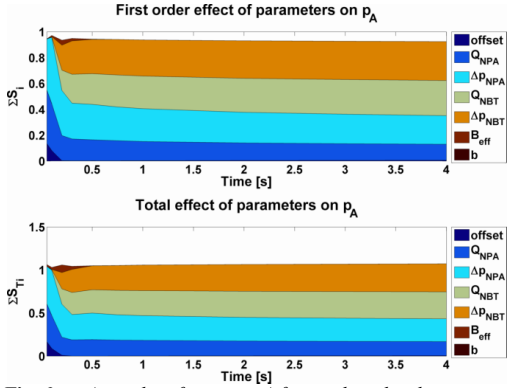


Fig. 6: *Area plot of pressure A first and total order sensitivity indices.*

The total order indices (Fig. 6:) drop similar to the first order indices. This indicates that the interactions between parameters are negligible. A difference between the first and total order indices is the influence of the somewhat larger effective bulk modulus.

Exact magnitudes are difficult to see from the area plot. Hence, the total order indices at selected time instants are gathered to Table 4:. We can see that effective bulk modulus B_{eff} has some impact on the system at time instants 0.20 and 0.30 seconds, even though its effect is smaller than the effects of nominal pressure differences and flow rates. This behaviour is expected because the system is in transient. What is interesting is that valve offset is very influential on pressure A few hundreds of a second into the experiment but loses its effect towards steady state. However, at no point is it more sensitive than the nominal parameters of notch PA.

Table 4: Pressure A total order sensitivity indices.

Time [s]	*offset*	$Q_{N,PA}$	$\Delta p_{N,PA}$	$Q_{N,BT}$	$\Delta p_{N,BT}$	B_{eff}	*b*
0.05	0.169	0.440	0.434	0.001	0.000	0.022	0.000
0.10	0.089	0.383	0.526	0.004	0.004	0.028	0.000
0.20	0.011	0.209	0.383	0.178	0.189	0.088	0.000
0.30	0.002	0.188	0.294	0.256	0.268	0.036	0.000
0.50	0.001	0.192	0.307	0.270	0.278	0.001	0.000
0.75	0.003	0.184	0.294	0.283	0.289	0.000	0.000
1.00	0.003	0.181	0.287	0.290	0.297	0.000	0.000
4.00	0.006	0.164	0.263	0.313	0.326	0.000	0.000

Other first and total order indices in transients are shown in Appendix 3. The pressure B first order indices behave opposite to pressure A indices. Specifically, the indices increase in the first few tenths of a second. The first order indices of velocity also behave differently; there is an increase in notch BT parameter indices and a decrease in notch PA parameter indices. The total order index of effective bulk modulus B_{eff} causes remarkable variance to pressure B, and the most variance to velocity at the beginning of the analysis. As time progresses, the effective bulk modulus loses its influence.

4.2 Sensitivity for Leakages

The analysis is carried out by studying the effects of internal leakage and external leakages in chambers A and B on pressures and velocity. The simultaneously varied parameters and their ranges are presented in Table 5:.

Table 5: Leakage sensitivity analysis parameters.

Parameter	Explanation	Lower bound [m^3/s $Pa^{-1/2}$]	Upper bound [m^3/s $Pa^{-1/2}$]
K_{int}	Internal leakage flow coefficient	0	$5.27*10^{-9}$
K_{extA}	External leakage A flow coefficient	0	$5.27*10^{-9}$
K_{extB}	External leakage B flow coefficient	0	$5.27*10^{-9}$

The upper bounds were chosen so that each leakage flow rate is 1 L/min with a pressure difference of 10 MPa, approximately 2.5 % of valve flow rate.

The leakage flows were modelled as turbulent and were added to the model at this stage. The flow equations (41)-(43) for internal leakage and external leakages A and B where the tank pressure is assumed to be zero are:

$$Q_{\text{int}} = K_{\text{int}}\sqrt{p_{\text{A}} - p_{\text{B}}} \tag{41}$$

$$Q_{\text{extA}} = K_{\text{extA}}\sqrt{p_{\text{A}}} \tag{42}$$

$$Q_{\text{extB}} = K_{\text{extB}}\sqrt{p_{\text{B}}} \tag{43}$$

Fig. 7: shows the first and total order effects of leakage faults on pressures and velocity in steady state. The first and total indices are approximately the same, the only difference being the effect of internal leakage on pressure A. This indicates only a minor amount of interactions between the leakage parameters K_{int}, K_{extA} and K_{extB}.

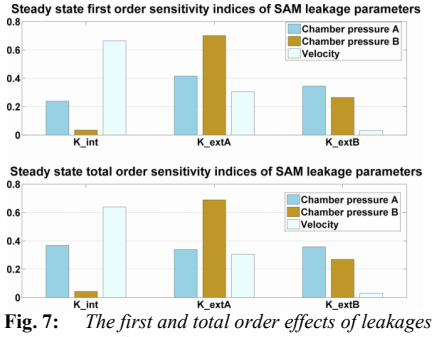


Fig. 7: *The first and total order effects of leakages in steady state.*

External leakage A causes remarkable variance to pressures A and B, and velocity. For explanation, consider the changes that occur as a consequence of external leakage A. When the leakage appears, it leads into a pressure drop in chamber A and a velocity decrease. As a consequence, the resistive pressure B drops. Pressure B is more sensitive to external leakage A than its own leakage because of the asymmetrical cylinder, the loading condition and the extending movement.

Consider the effects of external leakage B. That leakage reduces pressure B, which causes a mild increase in velocity (Fig. 7:). The influence on velocity is small, as external leakage B only lowers the motion-resistive pressure, and does not directly affect the driving pressure A. However, the results show that external leakage B, of course, (indirectly) affects pressure A through the motion equation. In this system, with its characteristics by the loading condition, the effect of external leakage B on pressure A was actually larger than on B.

Finally, look at the procedure when an internal leakage occurs. At first the internal leakage reduces pressure A and increases pressure B causing the velocity to decrease. However, the situation changes as time progresses since the increased flow into chamber B increases steady pressure B. Therefore, the pressure A also increases to balance. Finally, the velocity continues to drop, and internal leakage is clearly the most responsible for the variance in velocity (Fig. 7:). The influence of internal leakage on pressure B is minor, but could be larger in retraction. The leakage parameters are ranked to Table 6:.

Table 6: The ranking of leakage parameters according to their total order indices.

Ranking	p_{A}	p_{B}	v
1	K_{int}	K_{extA}	K_{int}
2	K_{extB}	K_{extB}	K_{extA}
3	K_{extA}	K_{int}	K_{extB}

The first order indices as a function of time for pressure A are presented in Fig. 8:. The external leakage A and internal leakage are the most influential leakages in the beginning, but as time progresses, and the flow-resistive pressure in chamber B develops, the external leakage B causes more and more variance to pressure A. At the same time the effect of external leakage A decreases and the effect of internal leakage increases.

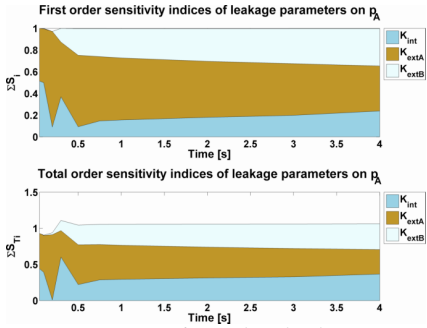


Fig. 8: *Pressure A first and total order sensitivity indices.*

The first order indices are shown with numerical figures in Table 7:.

Table 7: Pressure A first order sensitivity indices.

Time [s]	K_{int}	K_{extA}	K_{extB}
0.05	0.5141	0.4858	0.0000
0.10	0.4987	0.4994	0.0032
0.20	0.0908	0.8804	0.0000
0.30	0.3687	0.5042	0.1255
0.50	0.0928	0.6583	0.2454
0.75	0.1465	0.5928	0.2572
1.00	0.1552	0.5715	0.2695
4.00	0.2380	0.4145	0.3443

The total order indices are similar to the first order indices of K_{int} and K_{extB}, with the index of K_{extA} behaving differently. In the beginning, its influence increases as opposed to the decrease in first order indices, which flags interaction of external leakage A with other parameters. The total order indices are in Table 8:.

Table 8: Pressure A total order sensitivity indices.

Time [s]	K_{int}	K_{extA}	K_{extB}
0.05	0.4421	0.4872	0.000
0.10	0.3915	0.5107	0.0035
0.20	0.0098	0.9012	0.0264
0.30	0.6019	0.3660	0.1402
0.50	0.2219	0.5498	0.2718
0.75	0.2872	0.4882	0.2793
1.00	0.2944	0.4709	0.2906
4.00	0.3674	0.3385	0.3568

The first and total order sensitivity in transients for pressure B and velocity are in Appendix 4. In short, internal leakage causes the most variance to pressure B in the first tenth of a second; it is almost solely responsible for the variance. As time progresses, the influence of internal leakage decreases and the effects of external leakage A and B increase. The total order indices regarding pressure B and velocity are the same as the first order indices proving that there is no interaction between parameters. Throughout the analysis, the influence of external leakage B on velocity is nonexistent. As time progresses, internal leakage causes somewhat more variance to velocity, whereas external leakage A causes somewhat less.

5 Conclusions and Future Work

A generic Global Sensitivity Analysis-based approach that can be applied systematically to verify the model parameter sensitivities used for the model-based fault detection was presented in this paper. The GSA was applied to a valve-controlled asymmetrical hydraulic cylinder driving a 1-DOF manipulator joint to study its model parameter sensitivities. The studied parameters were the nominal flow rate and nominal pressure difference in the pressure and return notch of the valve, effective bulk modulus, valve spool offset and viscous friction coefficient. The sensitivity analysis was restricted to the extending motion.

Nominal flow rate and nominal pressure difference in the pressure notch of the valve were shown to be the most sensitive parameters to pressure or velocity responses regardless of whether the system was at steady state or transient. The second most sensitive parameters were the nominal flow rate and nominal pressure difference in the return notch. The effective bulk modulus was the third most sensitive parameter which was sensitive in transient pressure and velocity responses. The fourth most sensitive parameter was the valve offset which was sensitive in the steady state and transient velocity responses. The sensitivity of viscous friction was negligible throughout the analysis.

These results prove that flow coefficients should be identified to be as accurate as possible, since they had the largest sensitivity indices, and so the most effect on system outputs. Moreover, the identification of effective bulk modulus should be a second priority to facilitate model-based fault detection.

A leakage fault sensitivity analysis was also carried out to show the outputs from which the external leakage A, B and internal leakage could be best detected. The analysis proved that all leakage types can be detected with almost equal quality from the cylinder piston side pressure during transients or steady state during extension. From rod side pressure, all but the internal leakage in steady state and the external leakage B in transient are easily detectable. The rod side pressure was observed to be especially sensitive to internal leakage in transients. External leakage B was shown to be difficult to recognize from velocity in transients and steady state, so pressures are a prime candidate for detecting leakages.

The sensitivity indices can capture intuitively sensitive parameters and parameters whose sensitivity is more difficult to see. Whether the model is simple or complex, it is beneficial to systematically rank the parameters according to sensitivities since it decreases

the work needed in identifying parameters. The results of this part 1 will be used in part 2 where a scheme for detecting and isolating certain leakage and valve faults from a hydraulic system operating in various operating conditions is devised.

Acknowledgement

This work was funded by the Academy of Finland under the project 133273, Sensor network based intelligent condition monitoring of mobile machinery. The authors gratefully acknowledge the Academy of Finland for the financial support.

References

Archer, G., Saltelli, A. and **Sobol', I. M.** 1997. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. Statistical computation and simulation, volume 58, issue 2, 1997, pp. 99-120.

Beiner, L. and **Mattila, J.** 1999. An improved pseudoinverse solution for redundant hydraulic manipulators. Robotica, volume 17, pp. 173-179.

Canudas de Wit, C., Olsson, H., Åström, K. J. and **Lischinsky, P.** 1995. A new model for control of systems with friction. IEEE transactions on automatic control, volume 40, issue 3, pp. 419-425.

Hall, J. W., Boyce, S. A., Wang, Y., Dawson, R. J., Tarantola, S. and **Saltelli, A.** 2009. Sensitivity Analysis for Hydraulic Models. Journal of hydraulic engineering, November 2009.

Kaliatka, A., Kopustinskas, V. and **Vaišnoras, M.** 2009. Water hammer model sensitivity study by the FAST method. Energetika, volume 55, issue 1, pp. 13-19.

Kim, S. D., Cho, H. S. and **Lee, C. O.** 1987. A parameter sensitivity analysis for the dynamic model of a variable displacement axial piston pump. Proc. of IMechE, volume 201, issue C4.

Pastres, R., Franco, D., Pecenik, G., Solidoro, C. and **Dejak, C.** 1997. Local sensitivity analysis of a distributed parameters water quality model. Reliability engineering & system safety, volume 57, issue 1, July 1997, pp. 21-30.

Saltelli, A. 2002. Making best use of model evaluations to compute sensitivity indices. Computer physics communications, volume 145, issue 2, 15 May 2002, pp. 280-297.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and **Tarantola, S.** 2008. Global sensitivity analysis: the primer. John Wiley & Sons. 292 p.

Sobol', I.M. and **Kucherenko, S.S.** 2005. On global sensitivity analysis of quasi-Monte Carlo algorithms. Monte Carlo methods and applications, volume 11, issue 1, pp. 1-9.

Sobol', I. M., Tarantola, S., Gatelli, D., Kucherenko, S.S. and **Mauntz, W.** 2007. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. Reliability engineering & system safety, volume 92, issue 7, July 2007, pp. 957-960.

Watton, J. 1989. Fluid power systems: Modeling, simulation, analog and microcomputer control. Prentice Hall International (UK) Ltd. 490 p.

Appendix 1 - System Components

Table 9: The components in the SAM test bed.

Part	Model and specifications
Cylinder	Ø80/45-545
4/3-directional valve	Bosch Rexroth servo solenoid 4WRPEH 6 C3B24L-2X/G24K0/A1M (24 L/min @ 3.5 MPa)
Restrictor valve	Tognella needle valve FT257/2-38 (30 L/min @ 40 MPa)
Pressure transmitter	Trafag 8891.74 (0-25 MPa)
Pressure transmitter	Druck PTX 1400 (0-25 MPa)
Angle encoder	Heidenhain 376 886-0B (0.007 °/pulse)

Appendix 2 - Nomenclature and SAM Parameters

Parameter	Explanation	Value
A_A	Piston area	$\pi*(0.080)^2/4$ [m^2]
A_B	Piston rod area	$A_A - \pi*(0.045)^2/4$ [m^2]
b	Viscous friction	2500 Ns/m
d_r	Damping ratio	1
F_S	Static friction	4000 N
F_C	Coulomb's friction	1000 N
$\mathbf{J}(\theta)$	Moment of inertia matrix	-
K	Gain from control signal to spool position	0.1
$K_{vPA,leak}$	Leakage flow coeff. in notch PA	$1.9*10^{-12}$ $m^3/(sPa^{1/2})$
$K_{vBT,leak}$	Leakage flow coeff. in notch BT	$1.7*10^{-12}$ $m^3/(sPa^{1/2})$

$K_{vPB,leak}$, $K_{vAT,leak}$	Leakage flow coefficients in notches PB and AT	$1*10^{-12}$ $m^3/(sPa^{1/2})$
L	Load distance from the boom joint	1.9 m
L_B	Boom length	4.5 m
L_{i1}	Distance between upper cylinder joint and boom joint	0.30 m
L_{i2}	Distance between lower cylinder joint and boom joint	1.04 m
L_{i3}	Distance between lower and upper cylinder joints	0.84 m
m_B	Boom mass	297 kg
m_L	Left load mass	494 kg
m_R	Right load mass	0 kg
offset	Valve offset from center position	0
p_S	Supply pressure	10 MPa
$Q_{N,PA}$, $Q_{N,BT}$, $Q_{N,PB}$, $Q_{N,AT}$	Nominal flow rate in notch PA, BT, PB and AT	24 L/min
$\mathbf{R}(\theta)$	Torque arm matrix	-
r_B	Boom height	0.2 m
V_{0A}, V_{0B}	Volumes in A and B chambers	$2*10^{-4}$ m^3
v_s	Veloc. of min. frict.	0.01 m/s
x_{max}	Stroke	0.545 m
$\alpha_1 + \alpha_2$	See Fig. 3:	0.415 rad
$\Delta p_{N,PA}$, $\Delta p_{N,BT}$, $\Delta p_{N,PB}$, $\Delta p_{N,AT}$	Nominal pressure differences in notch PA, BT, PB and AT	3.5 MPa
θ	Joint angle	0.728 rad, cyl. retracted
σ_0	Friction coeff. 0	$4*10^6$ N/m
σ_1	Friction coeff. 1	$2*(\sigma_0*m_{redu})^{1/2}$
ω_n	Spool natural freq.	2*pi*20 rad/s

Appendix 3 - Sensitivity Indices of SAM

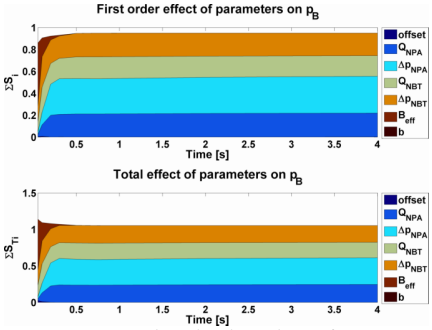


Fig. 9: *First and total order indices of parameters on pressure B.*

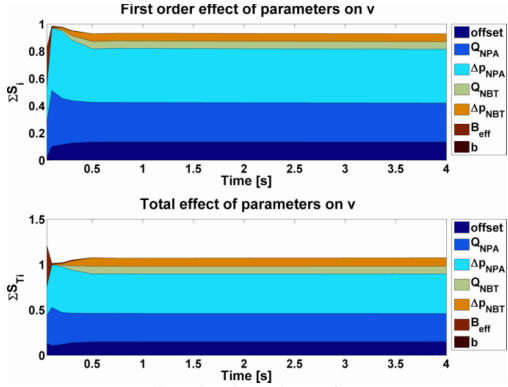


Fig. 10: *First and total order indices of parameters on velocity.*

Appendix 4 - Leakage Fault Sensitivity Indices of SAM

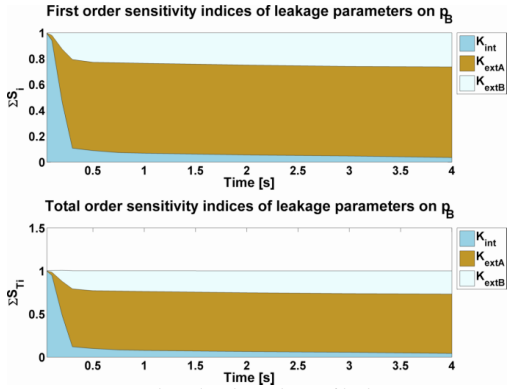


Fig. 11: *First and total order indices of leakage parameters on pressure B.*

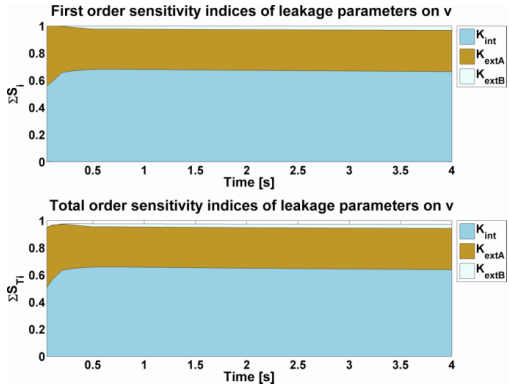


Fig. 12: *First and total order indices of leakage parameters on velocity.*

Publication II

Jarmo Nurmi, Jouni Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading condition, Part 2: Fault detection and isolation scheme," *International Journal of Fluid Power*, vol. 13, no. 1, pp. 17–27, March 2012.

This is an Accepted Manuscript of an article published by Taylor & Francis in International Journal of Fluid Power on 03/2012, available online: `http://www.tandfonline.com/10.1080/14399776.2012.10781043`

DETECTION AND ISOLATION OF LEAKAGE AND VALVE FAULTS IN HYDRAULIC SYSTEMS IN VARYING LOADING CONDITIONS, PART 2: FAULT DETECTION AND ISOLATION SCHEME

Jarmo Nurmi and Jouni Mattila

Tampere University of Technology, Department of Intelligent Hydraulics and Automation, P.O. Box 589, 33101 Tampere, Finland

E-mail: jarmo.nurmi@tut.fi

Abstract

Leakages and valve faults are among the most common faults in hydraulic systems. This paper studies the real-time detection and isolation of certain leakage and valve faults based on the results obtained in part one. In the first part, the mathematical model of a hydraulic test bed was analyzed with Global Sensitivity Analysis to facilitate a systematic and verified approach to model-based condition monitoring. In this paper, an Unscented Kalman Filter-based Fault Detection and Isolation scheme for leakage and valve faults of a generic servo valve-controlled hydraulic cylinder is devised. Compared to existing literature, the leakage and valve faults are decoupled from cylinder static and dynamic loading which makes the results generic and applicable to any servo valve-controlled hydraulic cylinder. Moreover, a more comprehensive set of fault patterns for the detection and isolation of leakages and valve faults with experimental and simulation results are presented. We show that detecting an external leakage of as small as 0.17 L/min is possible in some cases, but the accuracy of the method varies considerably. We also report why the isolation of valve faults from leakages is very difficult.

Keywords: fault detection and isolation, leakages, valve faults, varying load, unscented kalman filter, fault patterns

1 Introduction

The idea of model-based condition monitoring is to create a model output ŷ(k), which is subtracted from actual measurement y(k) to create a residual r(k) revealing the health of the system (Isermann, 2006). If the model is ideal, the residual remains at zero when the system is operating correctly. But when a fault is introduced, the residual deviates from zero, which is noticed by the fault detection process. Then the fault isolation process takes over and localizes the cause of the fault. The scheme as a whole is called Fault Detection and Isolation (FDI), see Fig. 1:.

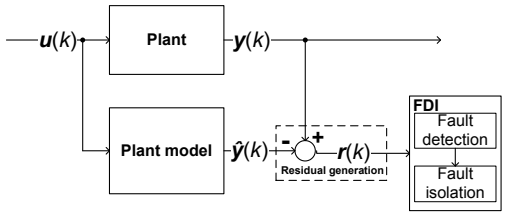


Fig. 1: *The model-based FDI scheme.*

In practice, measurements are noisy and perfect plant models are not possible. Therefore discrepancy between measured and modelled outputs is to be expected. For this reason state estimators (or Kalman Filters) which can consider modelling errors, measurement noise, and utilize measurements to correct model predictions are common in condition monitoring (An et al., 2008; Sepasi et al., 2010).

Previously, An and Sepehri (2008) proposed a method using a fault-free Extended Kalman Filter (EKF) to detect leakages with actuators under unknown external loading. Using the EKF to estimate the external force, they showed that external leakages out of the system and internal leakages across cylinder chambers as small as 0.25 L/min could be detected and isolated. Their approach was proven to work well with sinusoidal and fairly well with pseudorandom inputs.

More recently, Sepasi and Sassani (2010) applied the Unscented Kalman Filter (UKF) to detect leakages and load changes from a hydraulic system with a constant, known external force. They could detect and isolate leakage faults and load changes. However, results were provided using only sinusoidal inputs.

Chen (2010) devised a scheme to detect and isolate internal leakage and sensor offsets. The possibility of decoupling external force from state equations by considering velocity as an input was also proved.

Tan and Sepehri (2002) used the parameters of a nonlinear Volterra model to detect and isolate internal leakage, external leakage, incorrect supply pressure, and contamination in the fluid. Experimental results on the detection of incorrect supply pressure were shown, but the method was offline, which hampers its use for early fault detection. A similar issue affects the fault detection system by Le et al. (1998) where a neural network approach was shown to be sensitive to

relatively high leakages of over 1 L/min.

As opposed to model-based approaches, the use of the wavelet transform by Goharrizi et al. (2010a, 2010b) has produced good results by allowing the detection of an internal leakage of 0.124 L/min. But when external leakages were considered in Goharrizi et al. (2011), it was reported that external leakages of 0.30 L/min could be isolated from an internal leakage of 0.48 L/min, and furthermore external leakages cannot be localized to either side of the actuator, which has been proven to be possible with model-based approaches (An et al., 2008).

In this paper, we extend the methods of Sepasi and Sassani (2010), and An and Sepehri (2008) by treating a more extensive set of faults than those papers and adopting a similar method as Chen (2010) to obtain independence from varying load. The latter is possible as we have a sufficient quality position measurement from which we differentiate velocity, which eliminates the need to estimate them. Thus, we do not need to know the external force nor the load mass, as the information of the mechanism is included in the position and velocity measurement. Therefore, this scheme is more viable in generic hydraulic systems where the load can vary during operation. This paper utilizes the model and Global Sensitivity Analysis (GSA) that was presented for our test bed in (Nurmi and Mattila, 2011). This combines into a systematic approach to model-based condition monitoring compared to the ad hoc approaches currently present. An adaptive threshold is also proposed and experimental results are given with random control signals that are more plausible than sinusoidal inputs. The paper focuses on common leakage and valve faults (Watton, 2007).

This paper is organized as follows. In Section 2, the applicability of the method and the test bed are briefly discussed. In Section 3, the UKF algorithm is introduced and the reduced-order UKF is applied to the test bed. In Section 4, the capability of the UKF and the adaptive threshold scheme are experimentally tested in detecting and isolating leakages and with simulations in detecting and isolating valve faults.

2 Applicability of the Scheme and Test Bed

The FDI scheme used in this paper is applicable to a generic valve-controlled hydraulic cylinder that drives any of the n-DOF manipulator joints affected by any external force and inertia load (Fig. 2:). The scheme is considered to be suitable especially for detecting and isolating external and internal leakages.

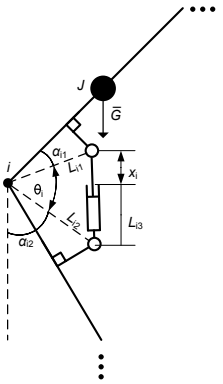


Fig. 2: *A manipulator joint driven by a hydraulic cylinder.*

2.1 Test Bed

As a case study to experimentally validate the scheme, the test bed, in Fig. 3: and Fig. 4:, is used in leakage fault study. It has a 4/3-directional valve that controls the joint cylinder, and three restrictor valves which emulate external leakages between the cylinder and the directional valve ('External leakage A' and 'External leakage B') and internal leakage across cylinder chambers ('Internal leakage'), for a list of system components see (Nurmi and Mattila 2011).

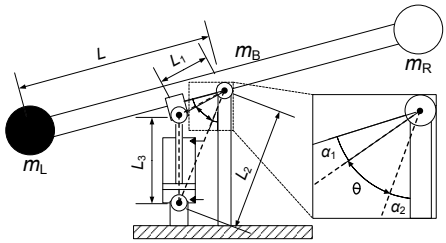


Fig. 3: *Illustration of the test bed.*

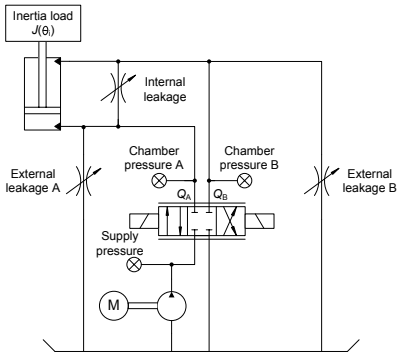


Fig. 4: *The hydraulic diagram of the test bed.*

In (Nurmi and Mattila, 2011), the model and GSA of the test bed were presented. In this paper, we use that model to simulate valve faults and utilize the GSA results in the verification of the UKF process model and in the development of the fault detection scheme.

3 Unscented Kalman Filter

In this section, an UKF scheme is devised to facilitate model-based FDI. The basis for the scheme originates from Sepasi and Sassani (2010) and An and Sepehri (2008). However, neither scheme is directly applicable to a generic hydraulic system where the load force and mass are not constant or known. Therefore, a modified version is used with decoupling of external force and load mass similar to (Chen, 2010).

This section is organized as follows. In Section 3.1, a generic discrete nonlinear system and its state estimation are introduced. Then in Section 3.2, the UKF algorithm is presented and implemented for the test bed in Section 3.3. Fault detection and isolation principles are discussed in Sections 3.4-3.5.

3.1 Discrete nonlinear system with noise and state estimation

The system is discrete with a nonlinear process $\boldsymbol{f}$ and measurement model $\boldsymbol{h}$ with noise vectors $\boldsymbol{w}$ and $\boldsymbol{v}$:

$$\begin{aligned} \boldsymbol{x}_{\mathrm{k+1}} &= \boldsymbol{f}(\boldsymbol{x}_{\mathrm{k}}, \boldsymbol{u}_{\mathrm{k}}, t_{\mathrm{k}}) + \boldsymbol{w}_{\mathrm{k}} \\ \boldsymbol{y}_{\mathrm{k+1}} &= \boldsymbol{h}(\boldsymbol{x}_{\mathrm{k+1}}, t_{\mathrm{k}}) + \boldsymbol{v}_{\mathrm{k+1}} \end{aligned} \tag{1}$$

where $\boldsymbol{x}$ is a N x 1 state vector in which N is the number of states, $\boldsymbol{u}$ is a Ux1 control vector in which U is the number of controls, t is the time, $\boldsymbol{w}$ is a Nx1 process noise vector, $\boldsymbol{y}$ is a Mx1 measurement vector in which M is the number of measurements, $\boldsymbol{v}$ is a Mx1 measurement noise vector and k is a prev. time instant.

The process noise $\boldsymbol{w}_{\mathrm{k}}$ and measurement noise $\boldsymbol{v}_{\mathrm{k+1}}$ are assumed to be Gaussian ($\mathcal{N}$), white (uncorrelated) and additive with zero mean and covariances $\mathbf{Q}_{\mathrm{k}}$ and $\mathbf{R}_{\mathrm{k+1}}$ with distributions:

$$\begin{aligned} \boldsymbol{w}_{\mathrm{k}} &\sim \mathcal{N}(0, \mathbf{Q}_{\mathrm{k}}) \\ \boldsymbol{v}_{\mathrm{k+1}} &\sim \mathcal{N}(0, \mathbf{R}_{\mathrm{k+1}}) \end{aligned} \tag{2}$$

Hydraulic measurements can be noisy, pressures especially. Considering the noise in the state estimator ensures that residuals are closer to zero in the fault-free situation, hence improving fault detection.

A Kalman-type state estimator for the nonlinear system in Eq. (1) is (Welch and Bishop, 2001):

$$\begin{aligned} \hat{\boldsymbol{x}}_{\mathrm{k+1}} &= \boldsymbol{f}(\hat{\boldsymbol{x}}_{\mathrm{k}}, \boldsymbol{u}_{\mathrm{k}}, t_{\mathrm{k}}) + \mathbf{K}_{\mathrm{k+1}}(\boldsymbol{y}_{\mathrm{k+1}} - \hat{\boldsymbol{y}}_{\mathrm{k+1}}) \\ \hat{\boldsymbol{y}}_{\mathrm{k+1}} &= \boldsymbol{h}(\hat{\boldsymbol{x}}_{\mathrm{k+1}}, t_{\mathrm{k}}) \end{aligned} \tag{3}$$

where $\hat{\boldsymbol{x}}$ is the state estimate vector of size (N - A) x 1 with the positive integer A denoting order-reduction. The innovation gain $\mathbf{K}_{\mathrm{k+1}}$ is chosen to minimize the mean squared error $\mathrm{E}[(\boldsymbol{x}_{\mathrm{k+1}} - \hat{\boldsymbol{x}}_{\mathrm{k+1}})^2]$. The optimal gain is derived in Simon (2006, pp. 318-320).

Nonlinear state estimation has no optimal solution since the innovation gain is dependent on covariances which are hard to accurately recover after the states are transformed through nonlinear functions. The non-optimal EKF circumvents the problem of nonlinearity by linearizing nonlinear functions around the previous states so that linear estimation techniques from the Kalman Filter (KF) can be applied. However, in the process it introduces approximation errors depending on the severity of the nonlinearity in functions $\boldsymbol{f}$ and $\boldsymbol{h}$.

An approach for tackling the problems of the EKF is the UKF, published by Julier et al. (1995). In (Julier and Uhlmann, 1997; Wan and van der Merwe, 2000) it is shown that the UKF approximates the true mean and covariance of the states more accurately than the EKF with Unscented Transformation (UT). The UT approximates the state distribution with deterministically chosen sigma points assuming that state variables are normally distributed.

Besides the accuracy advantage of UKF over EKF, UKF is also derivative-free, which is useful since calculating and writing long derivatives is error-prone. The given advantages motivate the choice of UKF.

3.2 Unscented Kalman Filter Algorithm

The recursive UKF algorithm can be described in a step by step manner as follows (Wan & van der Merwe 2000):

1. Initialize the filter, Eq. (4)
2. Estimate the a priori state vector $\hat{\boldsymbol{x}}_{k+1}^{-}$ (prediction)
   a. Generate sigma points around the previous estimate, Eq. (5)
   b. Propagate the sigma points through the nonlinear functions, Eq. (6)
   c. Calculate the state mean, Eq. (7)
3. Calculate the a priori error covariance $\mathbf{P}_{k+1}^{-}$, Eq. (8)
4. Estimate the a posteriori state vector $\hat{\boldsymbol{x}}_{k+1}$
   a. Unscented transformation for measurements (mean and covariance), Eq. (9)
   b. Calculate the cross-covariance between predicted states and measurements, Eq. (10)
   c. Calculate the Kalman gain, Eq. (11)
   d. Update state estimate, Eq. (12)
5. Calculate the a posteriori error covariance $\mathbf{P}_{k+1}$, Eq. (13)
6. Return to step 2

Step 1 is executed once and steps 2-6 are repeated. Steps 2 and 3 constitute the first UT, and step 4a the second. The steps correspond to the following equations:

$$\begin{aligned} \hat{\boldsymbol{x}}_0 &= \mathrm{E}(\boldsymbol{x}_0) \\ \mathbf{P}_0 &= \mathrm{E}[(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}}] \\ \mathrm{k} &= 0 \end{aligned} \tag{4}$$

$$\begin{aligned} \hat{\boldsymbol{x}}_{\mathrm{k}}^{(0)} &= \hat{\boldsymbol{x}}_k \\ \hat{\boldsymbol{x}}_{\mathrm{k}}^{(\mathrm{i})} &= \hat{\boldsymbol{x}}_{\mathrm{k}} + \tilde{\boldsymbol{x}}^{(\mathrm{i})}, \qquad \mathrm{i} = 1,2,3,\dots,2\mathrm{N} \\ \tilde{\boldsymbol{x}}^{(\mathrm{i})} &= \left(\sqrt{(\mathrm{L}+\lambda)\mathbf{P}_{\mathrm{k}}}\right)_{\mathrm{i}}^{\mathrm{T}}, \qquad \mathrm{i} = 1,2,3,\dots,\mathrm{N} \\ \tilde{\boldsymbol{x}}^{(\mathrm{i})} &= -\left(\sqrt{(\mathrm{L}+\lambda)\mathbf{P}_{\mathrm{k}}}\right)_{\mathrm{i}}^{\mathrm{T}}, \quad \mathrm{i} = \mathrm{N}+1,\dots,2\mathrm{N} \end{aligned} \tag{5}$$

$$\hat{\boldsymbol{x}}_{\mathrm{k+1}}^{(\mathrm{i})} = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{\mathrm{k}}^{(\mathrm{i})}, u_{\mathrm{k}}, t_{\mathrm{k}}\right) \tag{6}$$

$$\begin{aligned} w_{(\mathrm{mean})}^{(0)} &= \frac{\lambda}{\mathrm{L}+\lambda} \\ w_{(\mathrm{mean})}^{(\mathrm{i})} &= \frac{1}{2(\mathrm{L}+\lambda)}, \qquad \mathrm{i} = 1,2,3,\dots,2\mathrm{N} \\ \hat{\boldsymbol{x}}_{\mathrm{k+1}}^{-} &= \sum_{\mathrm{i}=0}^{2\mathrm{N}} w_{(\mathrm{mean})}^{(\mathrm{i})} \hat{\boldsymbol{x}}_{\mathrm{k+1}}^{(\mathrm{i})} \end{aligned} \tag{7}$$

$$\begin{aligned} w_{(\mathrm{cov})}^{(0)} &= \frac{\lambda}{\mathrm{L}+\lambda} + (1 - \alpha^2 + \beta) \\ w_{(\mathrm{cov})}^{(\mathrm{i})} &= \frac{1}{2(\mathrm{L}+\lambda)}, \qquad \mathrm{i} = 1,2,3,\dots,2\mathrm{N} \\ \mathbf{P}_{\mathrm{k+1}}^{-} &= \sum_{\mathrm{i}=0}^{2\mathrm{N}} w_{(\mathrm{cov})}^{(\mathrm{i})} \left(\hat{\boldsymbol{x}}_{\mathrm{k+1}}^{(\mathrm{i})} - \hat{\boldsymbol{x}}_{\mathrm{k+1}}^{-}\right)\left(\hat{\boldsymbol{x}}_{\mathrm{k+1}}^{(\mathrm{i})} - \hat{\boldsymbol{x}}_{\mathrm{k+1}}^{-}\right)^{\mathrm{T}} \\ &\quad + \mathbf{Q}_{\mathrm{k}} \end{aligned} \tag{8}$$

$$\begin{aligned} \hat{\boldsymbol{y}}_{\mathrm{k+1}}^{(\mathrm{i})} &= \boldsymbol{h}\left(\hat{\boldsymbol{x}}_{\mathrm{k+1}}^{(\mathrm{i})}, u_{\mathrm{k}}, t_{\mathrm{k}}\right) \\ \hat{\boldsymbol{y}}_{\mathrm{k+1}} &= \sum_{\mathrm{i}=0}^{2\mathrm{N}} w_{(\mathrm{mean})}^{(\mathrm{i})} \hat{\boldsymbol{y}}_{\mathrm{k+1}}^{(\mathrm{i})} \\ \mathbf{P}_{\mathrm{yy}} &= \sum_{\mathrm{i}=0}^{2\mathrm{N}} w_{(\mathrm{cov})}^{(\mathrm{i})} \left(\hat{\boldsymbol{y}}_{\mathrm{k+1}}^{(\mathrm{i})} - \hat{\boldsymbol{y}}_{\mathrm{k+1}}\right)\left(\hat{\boldsymbol{y}}_{\mathrm{k+1}}^{(\mathrm{i})} - \hat{\boldsymbol{y}}_{\mathrm{k+1}}\right)^{\mathrm{T}} \\ &\quad + R_{\mathrm{k+1}} \end{aligned} \tag{9}$$

$$\mathbf{P}_{\mathrm{xy}} = \sum_{\mathrm{i}=0}^{2\mathrm{N}} w_{(\mathrm{cov})}^{(i)} \left(\hat{\boldsymbol{x}}_{\mathrm{k+1}}^{(\mathrm{i})} - \hat{\boldsymbol{x}}_{\mathrm{k+1}}\right)\left(\hat{\boldsymbol{y}}_{\mathrm{k+1}}^{(\mathrm{i})} - \hat{\boldsymbol{y}}_{\mathrm{k+1}}\right)^{\mathrm{T}} \tag{10}$$

$$\mathbf{K}_{\mathrm{k+1}} = \mathbf{P}_{\mathrm{xy}} \mathbf{P}_{\mathrm{yy}}^{-1} \tag{11}$$

$$\hat{\boldsymbol{x}}_{\mathrm{k+1}} = \hat{\boldsymbol{x}}_{\mathrm{k+1}}^{-} + \mathbf{K}_{\mathrm{k+1}}(\boldsymbol{y}_{\mathrm{k+1}} - \hat{\boldsymbol{y}}_{\mathrm{k+1}}) \tag{12}$$

$$\begin{aligned} \boldsymbol{P}_{\mathrm{k+1}} &= \mathbf{P}_{\mathrm{k+1}}^{-} - \mathbf{K}_{\mathrm{k+1}} \mathbf{P}_{\mathrm{yy}} \mathbf{K}_{\mathrm{k+1}}^{\mathrm{T}} \\ &= \mathbf{P}_{\mathrm{k+1}}^{-} - \mathbf{P}_{\mathrm{xy}} \mathbf{P}_{\mathrm{yy}}^{-1} \mathbf{P}_{\mathrm{xy}}^{\mathrm{T}} \end{aligned} \tag{13}$$

where E is the expectation operator, w is a weighting coefficient, L is the dimension of the state vector and λ is a scaling parameter, satisfying $\lambda = \alpha^2(\mathrm{L} + \kappa) - \mathrm{L}$. The parameter α is a tuning factor which determines the spread of the sigma points. A typical value is 10^{-3}. The constant κ is a secondary tuning parameter. Usually it is chosen as zero. The constant β affects the weight of the first error covariance term. An optimal value is $\beta = 2$ for normally distributed states. The matrix square root in Eq. (5) should be calculated with Cholesky decomposition for computational efficiency.

If the measurement equations in function $\boldsymbol{h}$ are linear, steps 4a and 4b can be simplified. The equations in step 4a reduce to (Welch and Bishop, 2001):

$$\begin{aligned} \hat{\boldsymbol{y}}_{\mathrm{k+1}} &= \mathbf{H}\hat{\boldsymbol{x}}_{\mathrm{k+1}}^{-} \\ \boldsymbol{P}_{\mathrm{yy}} &= \mathbf{H}\mathbf{P}_{\mathrm{k+1}}^{-}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_{\mathrm{k+1}} \end{aligned} \tag{14}$$

Then step 4b reduces to:

$$\boldsymbol{P}_{\mathrm{xy}} = \boldsymbol{P}_{\mathrm{k+1}}^{-}\boldsymbol{H}^{\mathrm{T}} \tag{15}$$

where **H** is a measurement matrix of size M x N.

3.3 Unscented Kalman Filter implementation for the test bed

The online estimation of unknown load variables is possible (An et al., 2008), but not very feasible for FDI purposes because the UKF might compensate a fault by incorrectly estimating the load variables, hence making the fault undetectable. The problem is solved with the inclusion of position and velocity measurements to control vector $\boldsymbol{u}$ (Chen, 2010). The control vector then becomes:

$$\boldsymbol{u} = [x, \dot{x}, u_{\mathrm{c}}, p_{\mathrm{s}}] = [u_1, u_2, u_3, u_4] \tag{16}$$

where x is the position, $\dot{x}$ is the velocity, u_{c} is the valve control signal and p_{s} is the supply pressure. The position and velocity measurements could also be included to the state vector for filtering.

In the test bed, boom angle was measured and converted to piston position from which velocity was differentiated.

3.3.1 Process model

The task of the UKF is to estimate pressures p_{A} and p_{B}, spool position x_{s} and spool velocity $\dot{x}_{\mathrm{s}}$. The reduced-order state vector is thus:

$$\boldsymbol{x} = [p_{\mathrm{A}}, p_{\mathrm{B}}, x_{\mathrm{s}}, \dot{x}_{\mathrm{s}}]^{\mathrm{T}} = [x_1, x_2, x_3, x_4]^{\mathrm{T}} \tag{17}$$

Consequently, the discrete-time state space representation from (Nurmi and Mattila, 2011), Eq. (22), reduces to:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + T \begin{bmatrix} \frac{B_{\mathrm{effA}}}{A_{\mathrm{A}} u_1(k) + V_{0\mathrm{A}}} \left(Q_{\mathrm{A}}(x_1(k), x_3(k)) - u_2(k) A_{\mathrm{A}}\right) \\ \frac{B_{\mathrm{effB}}}{A_{\mathrm{A}}(x_{\max} - u_1(k)) + V_{0\mathrm{B}}} \left(Q_{\mathrm{B}}(x_2(k), x_3(k)) + u_2(k) A_{\mathrm{B}}\right) \\ x_4(k) \\ K\omega_{\mathrm{n}} u_3(k) - 2\omega_{\mathrm{n}} d_{\mathrm{r}} x_4(k) - \omega_{\mathrm{n}}^2 x_3(k) \end{bmatrix} \tag{18}$$

In the GSA in (Nurmi and Mattila, 2011), the

effective bulk moduli B_{effA} and B_{effB} were shown to be somewhat influential in transients, so effort was used to correctly identify them. They were found to be dependent on piston position. In particular, B_{effA} was quite small when the piston was completely retracted but gradually grew as the piston extended. The following equations taking the flexible volume of the hoses into consideration gave a good approximation:

$$B_{\mathrm{effA}} = \frac{B_{\mathrm{o}}B_{\mathrm{h}}(A_{\mathrm{A}}u_1(k) + V_{0\mathrm{A}})}{(A_{\mathrm{A}}u_1(k) + V_{0\mathrm{A}})B_{\mathrm{h}} + V_{\mathrm{h}}B_{\mathrm{o}}}$$
$$B_{\mathrm{effB}} = \frac{B_{\mathrm{o}}B_{\mathrm{h}}(A_{\mathrm{B}}(x_{\max} - u_1(k)) + V_{0\mathrm{B}})}{(A_{\mathrm{B}}(x_{\max} - u_1(k)) + V_{0\mathrm{B}})B_{\mathrm{h}} + V_{\mathrm{h}}B_{\mathrm{o}}} \quad (19)$$

where B_{o} is the bulk modulus of oil, B_{h} is the bulk modulus of the hose and V_{h} is the volume of the hose.

The flow coefficients, shown to be sensitive parameters and treated as constants in the GSA (Nurmi and Mattila, 2011), were not constants but nonlinear functions of spool position. To improve modelling accuracy each flow coefficient was fitted to a third-order polynomial (Muenchhof and Beck, 2008):

$$K_{\mathrm{v}}(x_3) = a_3x_3^3 + a_2x_3^2 + a_1x_3 + a_0 \quad (20)$$

Flow coefficient sample points were obtained offline by applying nonlinear parameter estimation techniques to step responses of the valve. Fitting a third-order polynomial to the sample points gave the best compromise between accuracy and complexity; see Root Mean Square Errors (RMSE) in Table 2:. The fitted polynomials are shown in Fig. 5: with the polynomial coefficients given in Table 1:.

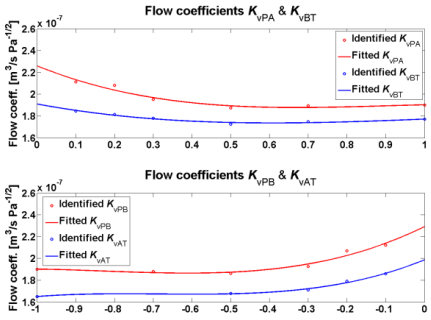


Fig. 5: *Flow coefficients of the valve.*

Table 1: Flow coefficient polynomials.

K_{vPA}	$(-5.121*10^{-8}x_3^3 + 1.556*10^{-7}x_3^2 - 1.377*10^{-7}x_3 + 2.256*10^{-7})*m^3/(s*Pa^{1/2})$
K_{vPB}	$(9.003*10^{-8}x_3^3 + 2.242*10^{-7}x_3^2 + 1.73*10^{-7}x_3 + 2.291*10^{-7})*m^3/(s*Pa^{1/2})$
K_{vAT}	$(1.031*10^{-7}x_3^3 + 2.118*10^{-7}x_3^2 + 1.423*10^{-7}x_3 + 1.986*10^{-7})*m^3/(s*Pa^{1/2})$
K_{vBT}	$(-2.371*10^{-8}x_3^3 + 7.573*10^{-8}x_3^2 - 6.588*10^{-8}x_3 + 1.91*10^{-7})*m^3/(s*Pa^{1/2})$

Table 2: A comparison of the goodness-of-fits between 1st and 3rd order polynomials.

Flow coefficient	Polynomial degree	RMSE [m^3/s Pa$^{-1/2}$]
K_{vPA}	1	$7.11*10^{-9}$
K_{vPA}	3	$\mathbf{3.94*10^{-9}}$
K_{vBT}	1	$3.76*10^{-9}$
K_{vBT}	3	$\mathbf{1.49*10^{-9}}$
K_{vPB}	1	$9.21*10^{-9}$
K_{vPB}	3	$\mathbf{4.96*10^{-9}}$
K_{vAT}	1	$5.38*10^{-9}$
K_{vAT}	3	$\mathbf{1.64*10^{-9}}$

In reality, the flow coefficients are also dependent on fluid temperature since the viscosity of the fluid changes with temperature. This modelling was omitted.

3.3.2 Initialization

The UKF is initialized as follows:

$$\begin{aligned} \boldsymbol{x}_0 &= [0,0,0,0]^{\mathrm{T}} \\ \boldsymbol{P}_0 &= \mathrm{diag}([10^{12}, 10^{12}, 10^{-3}, 10^{-3}]) \\ \boldsymbol{R} &= \mathrm{diag}([10^{10}, 10^{10}]) \\ \boldsymbol{Q} &= \mathrm{diag}([10^{8}, 10^{8}, 10^{-10}, 10^{-10}]) \\ \alpha &= 0.001 \\ \beta &= 2 \\ \kappa &= 0 \end{aligned} \quad (21)$$

where $\boldsymbol{x}_0$ is the initial state, $\mathbf{P}_0$ is the state covariance matrix, $\mathbf{R}$ is the measurement noise matrix and $\mathbf{Q}$ is the process noise matrix.

The tuning parameters α, β and κ were chosen according to existing literature (Wan and van der Merwe, 2000). $\mathbf{R}$ was chosen to represent measurement noise. The standard deviations of the pressure sensor readings were roughly 0.1 MPa.

The process noise covariance matrix $\mathbf{Q}$ represents modeling errors. It proved important to find a balance between process and measurement noise. The variances of process noise were chosen slightly smaller than the variances of measurement noise.

The measurement equations were linear, so the algorithm was reduced according to Eq. (14) and (15). The measurement matrix $\mathbf{H}$ was:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (22)$$

3.4 Fault detection principles

Pressures A and B residuals were calculated for detecting faults, which was justified on the basis of the GSA results provided in (Nurmi and Mattila, 2011). The residuals $r(k)$ were calculated as follows:

$$r(k) = p(k) - \hat{p}(k) \quad (23)$$

where $p(k)$ and $\hat{p}(k)$ are the measured and estimated

pressures, respectively, and k denotes the current time instant. The residuals were averaged within a moving 5-second window to remove the effect of brief estimation errors. The residual average $\mu_r(k)$ was calculated recursively with (Muenchhof and Isermann, 2005):

$$\mu_{\mathrm{r}}(k) = \mu_{\mathrm{r}}(k-1) + \frac{1}{N}[r(k) - r(k-N)] \qquad (24)$$

where N, the sample size, was 5000. The recursive formula, although computationally efficient, requires that N samples are stored in memory.

Because the process model cannot be tuned to perfection, a threshold is needed that ensures that the ratio between false alarms (false positives) and undetectable faults (false negatives) is as low as possible. To clarify, the threshold should be constructed in a way that that false alarms are minimized, but at the same time the threshold should as low as possible so that small faults can be detected. Of course to achieve this, the most important factor is the accuracy of the UKF process model.

It is usually enough to use a positive, constant threshold. In this case, the constant positive threshold was not sufficient. A negative threshold was needed so that negative residuals could be used in fault isolation. Also a constant threshold did not work, since the residuals in a faultless situation were larger in chamber A than in B. The reason for this was the load force causing a higher A than B pressure. Consequently, a pressure-dependent threshold was created. A smooth threshold was obtained by averaging both pressures within a moving 5-second window.

Through careful experiments, the following threshold polynomial produced the best results in terms of few false alarms and satisfactory fault detection:

$$t_{\mathrm{pos}}\big(\mu_{\mathrm{p}}(k)\big) = 0.012 + 4 * 10^{-4}\mu_{\mathrm{p}}^2 \qquad (25)$$

The unit of the threshold was MPa. The magnitude of the first term was based on the accuracy of the model. The second term ensured that the threshold increased at a suitable rate. The negative threshold was simply $t_{\mathrm{neg}}\big(\mu_{\mathrm{p}}(k)\big) = -t_{\mathrm{pos}}\big(\mu_{\mathrm{p}}(k)\big)$.

3.5 Fault isolation patterns

Leakage and valve faults were studied (Fig. 6:). The leakage faults were divided into 'External leakage in chamber A', 'External leakage in chamber B', 'External leakage in chambers A and B' and 'Internal leakage'. The valve faults were divided according to relative opening into 'Stuck to closed position', 'Too small an opening' and 'Too large an opening'.

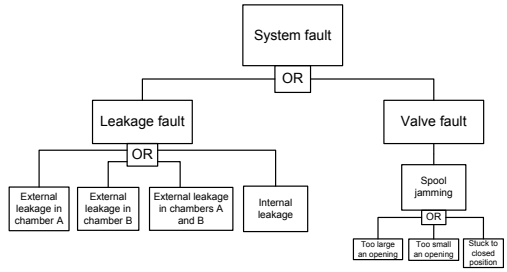


Fig. 6: *Fault tree (according to Isermann, 2006, p. 52).*

Once a fault was detected, the residuals and other variables were tested against fault patterns that were verified with experiments or simulations to isolate the fault:

Table 3: Fault patterns for leakage and valve faults.

#	Fault	$r_{p_A}\uparrow$	$r_{p_A}\downarrow$	$r_{p_B}\uparrow$	$r_{p_B}\downarrow$	$p_A > p_B$	Large r_{p_A} & r_{p_B}	$u_c > 0$
1	External leakage A	0	1	0	0	0 / 1	0	0 / 1
2	External leakage B	0	0	0	1	0 / 1	0	0 / 1
3	Internal leakage A→B	0	1	1	0	1	0	0 / 1
4	Internal leakage B→A	1	0	0	1	0	0	0 / 1
5	Spool jamming, too small an opening 1	0	1	1	0	0 / 1	0	1
6	Spool jamming, too small an opening 2	1	0	0	1	0 / 1	0	0
7	Spool jamming, too large an opening 1	1	0	0	1	0 / 1	0	1
8	Spool jamming, too large an opening 2	0	1	1	0	0 / 1	0	0
9	Spool jamming, stuck to closed position 1	0	1	1	0	0 / 1	1	1
10	Spool jamming, stuck to closed position 2	1	0	0	1	0 / 1	1	0

where ↑ denotes the crossing of the positive threshold, ↓ the crossing of the negative threshold, $r_{\mathrm{p_A}}$ is the pressure A residual, $r_{\mathrm{p_B}}$ is the pressure B residual and u_c is the valve control signal.

For example, an external leakage A causes the pressure A residual to cross the negative threshold ($r_{\mathrm{p_A}}\downarrow = 1$), and the residual $r_{\mathrm{p_B}}$ to remain within thresholds ($r_{\mathrm{p_B}}\uparrow = 0$ and $r_{\mathrm{p_B}}\downarrow = 0$). Considering the direction of the residual we could distinguish simultaneous external leakage A and B from internal leakage and from certain valve faults.

The faults where the valve opens too wide or too little can be *instantly* isolated from internal leakage with a 50-percent probability when both are considered as likely. The possibility of instant detection depends on the test '$p_A > p_B$'. For example if the first four binaries of the fault code are 0110, and the fifth, the

test '$p_A > p_B$' is false, there is offset in the spool position, so either the valve opened too little or too wide. If the test '$u_c > 0$' is true, the valve opening was too small. If it is false, the valve opening was too large. It is possible to isolate the valve fault when the control signal changes from positive to negative or vice versa by observing whether the residuals cross the opposite thresholds. Internal leakage, on the other hand, is not dependent on the sign of valve control signal.

The rationales behind the patterns are as follows. Consider external leakage A as an example. The GSA (Nurmi and Mattila 2011) proved that both pressures are sensitive to a chamber A leakage. However, only the residual r_{p_A} crosses the negative threshold, since the pressure differential $\dot{p}_A$ is missing a leakage flow term. The velocity also changes, as shown in the GSA (Nurmi and Mattila 2011). However, its effect to pressure residuals is minor. A similar description applies to external leakage B and internal leakage. However, in internal leakage faults a leakage is present in both chambers. In one chamber the leakage flow is negative, and in the other it is positive.

When the valve is given a positive control signal and it fails to open as much as it should (fault #5), the flow rate to chamber A is too small compared to a fault-free situation. Thus the pressure A measurement is smaller than the UKF estimate and consequently the pressure A residual crosses the negative threshold. At the same time, the pressure B residual crosses the positive threshold because the measured B pressure, as a consequence of the restricting action of the smaller notch BT opening, is larger than the estimated B pressure. Similar explanations apply to faults #6-10. In faults #9-10, the valve is completely closed, so the magnitudes of the residuals reveal the cause.

4 Results

Experimental results for detecting and isolating leakages are given in Section 4.1. In Section 4.2, the detection and isolation of valve faults is studied with simulations.

4.1 Experimental results

The experimental results consist of external leakage A, external leakage B, simultaneous external leakage A and B, and internal leakage. The valve was controlled with fairly random control signals (Fig. 7:).

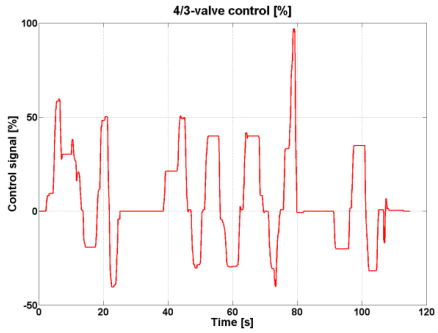


Fig. 7: *Example control signal to the 4/3-directional valve that was used in the external leakage B experiment.*

In the residual figures, the black vertical line shows the time when the fault was added, the solid red line the residuals, and the dashed black lines the thresholds.

4.1.1 External leakage in chamber A

The external leakage A was added to the system at around the 35th second. The evolution of pressure A and pressure B measurements and estimates are given in Fig. 8:. The estimates are in blue and the measurements in red colour.

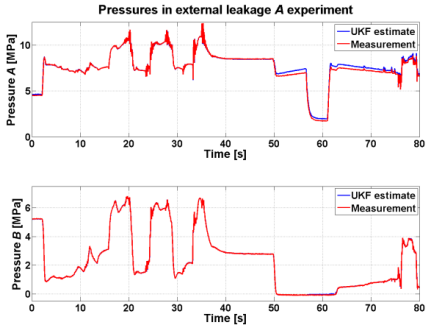


Fig. 8: *The evolution of pressure A and B estimate and measurement.*

Between the 35th and 50th second the difference between the pressure A measurement and its estimate is not clear-cut. From the 50th second onwards, the difference becomes clear and is indistinguishable between pressure B measurements and estimates (Fig. 8:).

The external leakage A was detectable two seconds later when the pressure A crossed the negative threshold at -0.04 MPa, see Fig. 9:. Therefore, according to the fault patterns in Table 3:, the fault could be isolated as external leakage A (fault #1). The magnitude of the residual indicates that the fault is severe. According to flow measurements, a detectable external leakage was close to 0.40 L/min, or 10 % of the flow rate passing the valve.

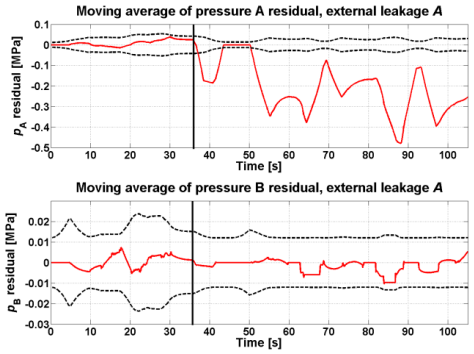


Fig. 9: *Pressure residuals (in solid red) with a varying external leakage in chamber A of average 1.40 L/min. The thresholds are the dashed black lines.*

The residuals were momentarily decreased to zero since there were some non-fault related discrepancies between measurements and UKF estimates when the pressure B was close to zero or the boom angle was zero. Excluding these situations, the scheme worked.

4.1.2 External leakage in chamber B

The external leakage B was added to the test bed at around the 30th second. The fault was detectable a few seconds later when pressure B crossed the negative threshold, as illustrated in Fig. 10:.

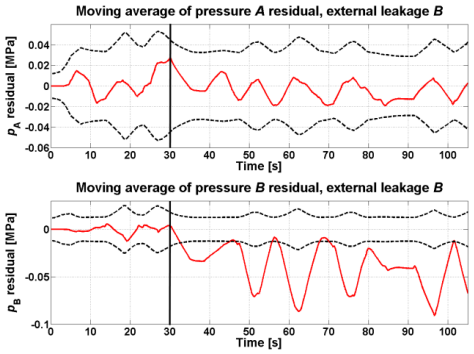


Fig. 10: *Pressure residuals with an external leakage in chamber B with average 0.50 L/min.*

According to fault patterns in Table 3:, the fault could be isolated as external leakage B (fault #2). As pressure B was smaller than pressure A, the thresholds for B residuals could be considerably smaller allowing for a smaller leakage to be detected. At the time of detection, the threshold was -0.015 MPa and the minimum detectable leakage approximately 0.17 L/min, or 5 % of the flow through the 4/3 directional valve, significantly smaller than the detectable leakage from chamber A. The leakage varied between 0.14 L/min and 1 L/min (see the varying magnitude of residual), but on average it was 0.50 L/min.

When the pressure A residual approached the threshold, the threshold increased, proving that the thresholds were indeed pressure dependent and so the proposed adaptive threshold worked.

4.1.3 Simultaneous external leakage in chambers A and B

An external leakage A and B were simultaneously introduced to the test bed at the 18th second, as shown in Fig. 11:. The external leakage A was detectable only a second later, but the external leakage B took over 30 seconds to detect. The reason for the slow detection was the decreased pressure A that decreased pressure B causing a minor leakage from chamber B. The leakage in chamber B rose to 0.18 L/min before the actual detection of the fault, but a short leakage peak of this magnitude could not be detected. The leakage peaked at 0.52 L/min (average 0.30 L/min) and 0.60 L/min (average 0.34 L/min) at 50 and 65 seconds, and at those instants the threshold of residual B was clearly crossed.

An external leakage A of 0.50 L/min could be detected as that was the leakage magnitude at the time of detection. The leakage averaged at 1.6 L/min between 40 and 70 seconds, but the residual during this period was well over the threshold.

The isolation follows the patterns in Table 3:. For the reasons in Section 4.1.1, the residuals were momentarily forced to zero.

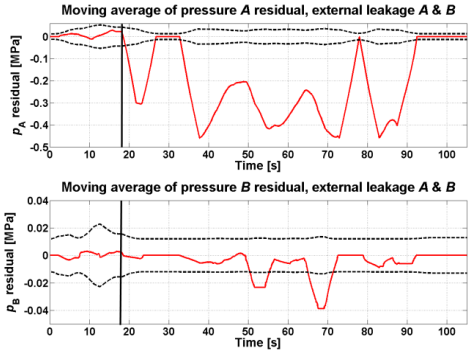


Fig. 11: *Pressure residuals with an external leakage in chambers A and B. An external leakage A of 11 % (average 1.60 L/min) and an external leakage B (average 0.30 L/min) of 5 % of flow through the valve were detectable.*

4.1.4 Internal leakage

The internal leakage was added to the test bed at the 30th second, as shown in Fig. 12:. The positive threshold of pressure B residual was crossed roughly three seconds sooner than the negative of pressure A residual. The leakage varied between 0.35 L/min and 2 L/min, and on average it was 0.94 L/min. During the experiment when the leakage dropped significantly

below the average, the thresholds remained in the fault range, showing that an internal leakage of below 0.50 L/min could be detected, or a leakage in the range of 5-10 % of the flow passing the 4/3-directional valve.

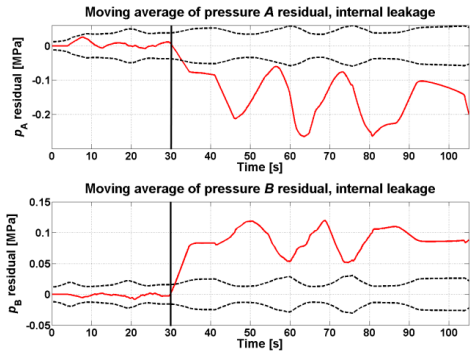


Fig. 12: *Pressure residuals with an internal leakage of average 0.94 L/min.*

The isolation of the internal leakage was not entirely straightforward following Table 3:. Since $p_A > p_B$ was always true, at the time of detection the fault was either an internal leakage (fault #3), or the spool had opened too wide (fault #7 and #8) or too little (fault #5 and #6). As time progressed, the residuals did not cross the other threshold as the control signal changed. Therefore the fault could be isolated as internal leakage.

4.2 Simulation results

The accuracy of the UKF process model in simulations guaranteed that the residuals stayed close to zero in a faultless situation. For consistency the adaptive threshold was used in simulations. White, normally distributed noise was added to simulation pressures so that they bore more of a resemblance to experimental measurements. In addition, the UKF parameters were retuned. The process noise variances of pressures were reduced to 10^4 Pa^2 and the measurement noise variances to 10^5 Pa^2. The 4/3-directional valve control signal used in simulations is shown in Fig. 13:.

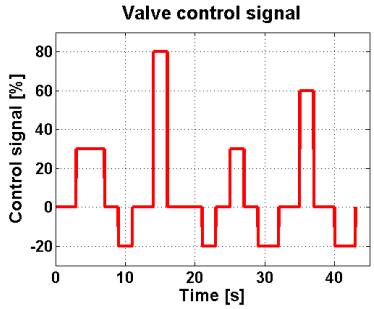


Fig. 13: *The valve control signal in simulation experiments.*

4.2.1 Too small valve opening

The 'too small valve opening' fault was introduced to the system at the 35th second, as is illustrated in Fig. 14:.

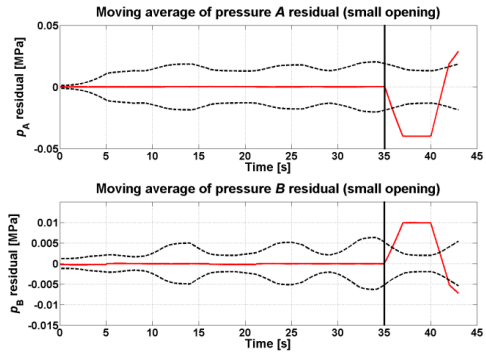


Fig. 14: *The spool jams so that the opening is too small.*

After a second, the fault was detected but could not be immediately isolated since pattern was similar to internal leakage (#3). Once both residuals, as a consequence of the valve control signal change, crossed the other threshold at approximately the 42th second, the fault could be isolated as too small valve opening fault (#5-6).

4.2.2 Too large valve opening

The fault 'too large valve opening' (#7-8) was added to the system at the 35th second, as shown in Fig. 15:. The residual behaviour was reversed compared to fault case 'too small valve opening'. Hence the isolation was possible immediately after detection.

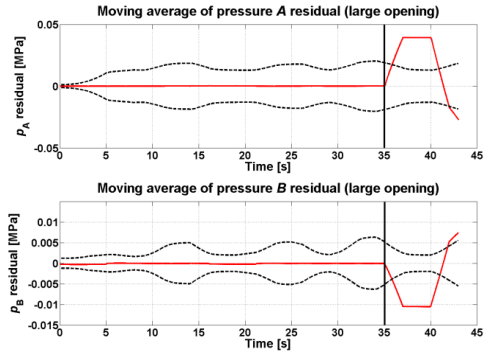


Fig. 15: *The spool jams so that the opening is too large.*

If the internal leakage would have occurred from chamber B to A, this fault could have been isolated, although not instantly following detection.

4.2.3 Valve gets stuck to closed position

At the 35th second, the valve spool got stuck to closed position, shown in Fig. 16:.

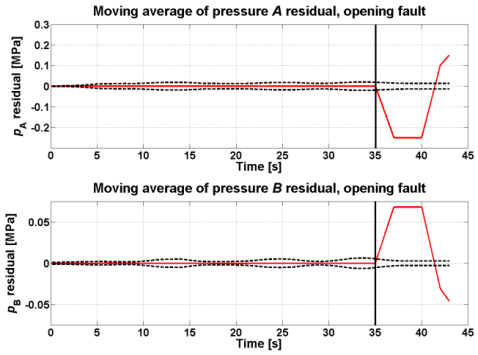


Fig. 16: *The valve spool stuck to closed position.*

The residual behaviour was similar to internal leakage (#3), except that the residuals crossed the opposite threshold when the valve control signal was reversed at the 40th second. The magnitudes of the residuals revealed the cause as being a closed valve (#9-10).

5 Conclusions and Future Work

In this paper a real-time scheme based on a reduced-order Unscented Kalman Filter (UKF) for detecting and isolating leakage and valve faults from a generic valve-controlled hydraulic cylinder driving a manipulator joint in varying loading conditions was devised and applied to a hydraulic boom test bed. The method is a practical load-independent solution for detecting and isolating especially leakages. In the paper, a comprehensive set of fault patterns were presented, and an adaptive threshold facilitating fault detection was devised. The basis for this work was founded in (Nurmi and Mattila, 2011), where a Global Sensitivity Analysis (GSA) of the test bed was carried out. The results of the analysis were used in this paper, showing that GSA facilitates a systematic and verified approach to model-based condition monitoring. The usefulness of GSA increases with more complicated nonlinear models.

The fault patterns were verified with simulation and experimental studies. The studies, with leakage patterns verified experimentally, showed the possibility of distinguishing external leakage A and B, and simultaneous external leakage A and B from valve faults (spool jamming to too large an opening, to too small an opening and stuck to closed position).

The lowest detectable external leakage was 0.17 L/min, but it varied between experiments. The pressure residuals alone were not enough to distinguish internal leakage and valve faults, so information from the sign of the control signal and the larger chamber pressure were used. The control signal test meant that the time from fault detection to isolation was considerably long in some valve opening or internal leakage fault cases. The fault patterns for those two different faults were found similar, and hence very difficult to distinguish.

The fault-to-isolation time could be shortened and the fault patterns could be expanded by utilizing a spool position measurement. This measurement, however, is not usually available. Moreover, the scheme already requires multiple measurements, and more measurements would increase the probability of sensor failures. The required measurements, however, are: cylinder chamber pressures A and B, supply pressure, valve control signal, piston position or boom angle, and piston velocity or boom angular velocity measurement. If a separate velocity measurement is not available, the velocity could be differentiated from position.

The scheme will be extended to mobile valves using position and velocity sensors more suitable for application domain specific environmental conditions in the future.

Acknowledgement

This work was funded by the Academy of Finland under the project 133273, Sensor network based intelligent condition monitoring of mobile machinery. The authors gratefully acknowledge the Academy of Finland for the financial support.

Nomenclature

μ_p	Pressure average	[MPa]
r	Residual	[MPa]
$t_{\mathrm{pos,neg}}(\mu_\mathrm{p})$	Positive and negative threshold polynomial	[MPa]
$\mathbf{H}$	Measurement matrix	[-]
$\mathbf{K}$	Kalman gain	[-]
$\mathbf{P}$	Posteriori state covariance matrix	[-]
$\mathbf{P}^-$	Priori state covariance matrix	[-]
$\mathbf{P}_{\mathrm{xy}}$	Cross-covariance matrix	[-]
$\mathbf{P}_{\mathrm{yy}}$	Measurement covariance matrix	[-]
$\mathbf{Q}$	Process noise covariance matrix	[-]
$\mathbf{R}$	Measurement noise covariance matrix	[-]
$\boldsymbol{x}$	State vector	[-]
$\hat{\boldsymbol{x}}$	State estimate vector	[-]
$\tilde{\boldsymbol{x}}^{(\mathrm{i})}$	Sigma point vector	[-]
$\boldsymbol{y}$	Measurement vector	[-]

$\boldsymbol{v}$	Measurement noise vector	[-]
$\boldsymbol{w}$	Process noise vector	[-]

Subscripts:

k	Discrete time instant

References

An, L. and **Sepehri, N.** 2008. Leakage fault detection in hydraulic actuators subject to unknown external loading. International journal of fluid power, volume 9, issue 2, August 2008, pp 15-25.

Chen, L. 2010. Model-based fault diagnosis and fault-tolerant control for a nonlinear electro-hydraulic system. Ph.D. Thesis. University of Kaiserslautern.

Goharrizi, A.Y., **Sepehri, N.** 2010a. A wavelet-based approach to internal seal damage diagnosis in hydraulic actuators. IEEE Trans. on Ind. Electron., vol. 57, no. 5, pp.1755-1762.

Goharrizi, A.Y., **Sepehri, N.** and **Wu, Y.** 2010b. A wavelet based approach for diagnosis of internal leakage in hydraulic actuators using on-line measurements. Int. Journal of Fluid Power, vol. 11, no.1, pp. 61-69.

Goharrizi, A.Y., **Sepehri, N.** and **Wu, Y.** 2011. A wavelet-based approach for online external leakage diagnosis and isolation from internal leakage in hydraulic actuators. Int. Journal of Fluid Power, vol. 12, no. 2, pp. 37-47.

Isermann, R. 2006. Fault-diagnosis systems, an introduction from fault detection to fault tolerance. Springer-Verlag Berlin Heidelberg. 475 p.

Julier, S. J., Uhlmann, J. K. and **Durrant-Whyte, H. F.** 1995. A new approach for filtering nonlinear systems. Proceedings of the American control conference, Seattle, Washington, 1995, pp. 1628–1632.

Julier, S. J. and **Uhlmann, J. K.** 1997. A new extension of the Kalman filter to nonlinear systems. Proceedings of AeroSense, the 11th international symposium on aerospace/defence sensing, simulation and controls, pp. 182-193.

Le, T. T., Watton, J. and **Pham, D. T.** 1997. Fault classification of fluid power systems using a dynamics feature extraction technique and neural networks, Proc. Instn. Mech. Engrs, Vol. 212, Part I, pp. 87-96.

Muenchhof, M. and **Isermann, R.** 2005. Comparison of change detection methods for a residual of a hydraulic servo axis. Proceedings of the 16th IFAC world congress, international federation of automatic control, Czech Republic, July 4-8, 2005.

Muenchhof, M. and **Beck, M.** 2008. Model adjustment and multi-model based fault diagnosis for hydraulic servo axis. Proceedings of the 17th world congress, the international federation of automatic control, Seoul, Korea, July 6-11, 2008.

Nurmi, J. and **Mattila, J.** 2011. Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 1: Global Sensitivity Analysis. International Journal of Fluid Power, volume 12, no. 3.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and **Tarantola, S.** 2008. Global sensitivity analysis: the primer. John Wiley & Sons. 292 p.

Sepasi, M. and **Sassani, F.** 2010. On-line fault diagnosis of hydraulic systems using unscented Kalman filter. International journal of control, automation, and systems, volume 8, issue 1, 2010, pp. 149-156.

Simon, D. 2006. Optimal state estimation: Kalman, H∞, and nonlinear approaches. Wiley-Interscience. 526 p.

Tan, H. and **Sepehri, N.** 2002. Parametric fault diagnosis for electrohydraulic cylinder drive units. IEEE Trans. Industrial Electronics, Vol. 49, pp. 96-106.

Wan, E. A. and **van der Merwe, R**. 2000. The Unscented Kalman Filter for nonlinear estimation. Proceedings of IEEE adaptive systems signal processing, communication and control symposium, October 2000, pp. 153–158.

Watton, J. 2007. Modelling, monitoring and diagnostic techniques for fluid power systems. Springer-Verlag London Ltd. 360 p.

Welch, G. and **Bishop, G.** 2001. An introduction to the Kalman filter. Department of computer science, University of North Carolina, Chapel Hill, TR95-041.

Publication III

Jarmo Nurmi, Janne Honkakorpi, Juho Vihonen, Jouni Mattila, "Micro-electromechanical system sensors in unscented Kalman filter-based condition monitoring of hydraulic systems," *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, Wollongong, Australia, July 9 – 12, 2013, pp. 354 – 361, http://dx.doi.org/10.1109/AIM.2013.6584117.

Micro-electromechanical System Sensors in Unscented Kalman Filter-based Condition Monitoring of Hydraulic Systems

Jarmo Nurmi, Janne Honkakorpi, Juho Vihonen, and Jouni Mattila

***Abstract*—A condition monitoring system of mechatronic design using an easy-to-install micro-electromechanical system-based (MEMS) motion sensor is developed in this paper, since the joint encoders and resolvers that require mechanical joint modifications are not often a realistic option for mobile machines. This paper presents an unscented Kalman filter-based (UKF) condition monitoring scheme for leakage detection in hydraulic actuator systems using the joint sensor feedback. By comparing the UKF residuals of the developed MEMS sensor to the residuals of a highly accurate reference encoder, we show that the developed easy-to-install and low-cost MEMS sensor is suitable for this condition monitoring task.**

I. Introduction

Hydraulically driven heavy-duty manipulators and cranes are widely used e.g. in material and cargo handling, construction and mining industries. These systems are still predominantly open-loop controlled by human operators and thus do not have any joint sensor instrumentation. However, the manufacturers of these machines have great interest to increase their after-sales market offering in particular in condition monitoring and machine operation optimization. Introducing advanced software-based functionality such as condition monitoring still requires motion sensors that would to be easily retrofitted into these machines. The traditional option for retrofitting would be joint angle encoders or resolvers, but these sensors are often of high-cost (> 300 €), not suitable for rugged environments and their installation requires modifications prone to mechanical failures. Therefore, we foresee that easy to retrofit, low-cost (~ 100€ mass-produced) MEMS sensor technology embedded with advanced signal processing is required to enable the conditions monitoring applications.

Previous work in the development of MEMS sensor technology includes e.g. [1], [2], [3] where measurements from multiple accelerometers were combined, and [4] where low-cost tilt sensor, rate gyro and accelerometer measurements were fused. In terms of condition monitoring, the use of similar MEMS sensor measurements has been limited to vibration analysis, see e.g. [5], [6]. In this paper, a different approach to condition monitoring is taken by considering the utilization of the developed MEMS motion sensor in a model-based load-independent condition monitoring algorithm for a hydraulic manipulator where from the two-fold mechatronic perspective, firstly the algorithm development for the robust, low-cost and easy-to-install embedded motion sensor prototype is presented, and secondly its application to condition monitoring is demonstrated. The algorithm is founded on the principle of complementary filtering to improve the quality of position and velocity measurements. The condition monitoring capability is demonstrated by comparison with a high accuracy reference sensor (joint angle encoder) in a generic and easily parametrisable scheme for a hydraulic application where for example a joint angle encoder is not a realistic design option to satisfy the requirements of easy installation and retrofitting to existing machines. The MEMS sensor on the other hand satisfies these requirements and is thus a suitable candidate to enable advanced functionality such as condition monitoring.

In reference to prior work by the authors, contrary to [7], the easily applicable MEMS-based sensor is here used in place of the unrealistic option, the joint angle encoder in unscented Kalman filter-based (UKF) fault detection. The UKF fault detection is based on the residual between the system model and measurement. Implementation of the load-independent UKF condition monitoring system with the MEMS sensor is novel in reference to prior work. Additionally, previously in [8] the successful application of the MEMS sensor to hydraulic manipulator state feedback control was shown, and so this paper expands the possible application range of the MEMS sensor to make it more attractive to mobile machine manufacturers looking to compete in the global market with novel functionalities and services.

This paper is organized as follows. In Section II, sensors in hydraulic manipulator control are discussed with a focus on the developed MEMS sensor prototype and its algorithms. The UKF and the reduced-order model of a servo valve-controlled hydraulic cylinder are introduced in Section III. In Section IV, the performance of the MEMS sensor is experimentally assessed and compared against encoder measurements. Section V discusses the obtained results. Finally, the conclusions are drawn in Section VI. The paper concludes with an appendix containing the UKF algorithm.

II. Sensors in Hydraulic Control

In hydraulic control applications, the design of a high-quality controller is of great importance in achieving desirable properties, such as accurate tracking and minimization of positioning errors. Regardless of controller design, these properties cannot be achieved without a high-

This work was supported in part by the Academy of Finland under project 133273.

Jarmo Nurmi and Janne Honkakorpi are researchers at the Dept. of Intelligent Hydraulics and Automation, Tampere University of Technology, e-mails: jarmo.nurmi@tut.fi, janne.honkakorpi@tut.fi.

Juho Vihonen is a post-doctoral researcher at the Dept. of Signal Processing, Tampere University of Technology, e-mail: juho.vihonen@tut.fi

Jouni Mattila is a professor of machine automation at the Dept. of Intelligent Hydraulics and Automation, Tampere University of Technology, e-mail: jouni.mattila@tut.fi.

quality measurement of the joint angle. Most often the joint angle is measured with a rotary resolver or encoder due to their accuracy. However, they are usually expensive and cannot be easily retrofitted to existing machines because of the requirement of contact to the joint axle or a shaft between adjacent links if the axle is non-rotary (Fig. 1). The installation of the encoder also requires high co-axial precision and if the installation is done poorly, the accuracy and mechanical durability of the sensor could be compromised.

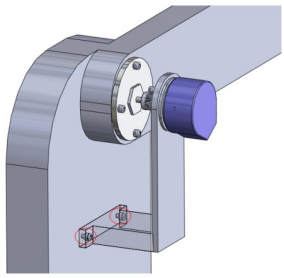


Figure 1. Installation of an angle encoder to the joint of a manipulator.

To overcome these limitations of the resolvers and encoders, the Department of Intelligent Hydraulics and Automation of Tampere University of Technology has developed a MEMS sensor module suitable for the harsh working conditions of mobile working machines where the MEMS chip integrates micromechanical structures and some signal conditioning electronics onto a single silicon chip. The mechanism sizes are minute and hence well-suited to cramped spaces with requirements of high chip-level integration. The module is encased in an epoxy-filled dust- and waterproof aluminium enclosure as shown in Fig. 2. The MEMS chip is the model SCC1300-D02 by Murata Electronics, which combines a three-axis accelerometer with a one-axis rate gyroscope.

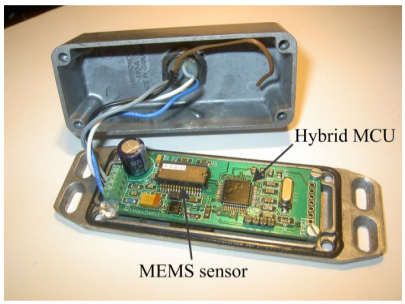


Figure 2. The developed MEMS sensor prototype (without epoxy filling).

The accelerometer measures linear acceleration along the three axes and can also measure inclination angles with the application of inverse trigonometric functions to observed vector components of gravitational acceleration. The digital output of the accelerometer has a resolution of 0.56 mg per least significant bit (LSB) which translates into a best-case inclination resolution of 0.032 deg when the accelerometer axis is parallel to ground. However, if the accelerometer is placed far away from the axis of rotation, the gravity sensing can be impaired. The digital signal of the gyro has a resolution of 0.02 deg/s per LSB. The MEMS chip has been combined with a 16-bit hybrid signal microcontroller unit (MCU) and a line driver chip for implementing low-level data operations and CANopen communication. This makes the sensor module well suited for retrofit-type integration into existing platforms.

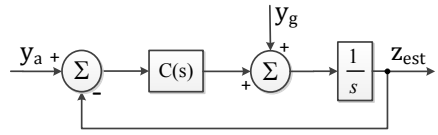


Figure 3. Complementary filter structure.

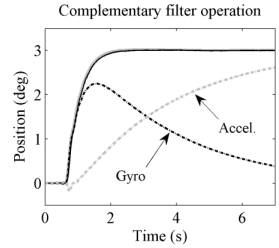


Figure 4. Example complementary filter output (solid black line) versus encoder position (solid grey line) when given a step input.

As stated previously, a recently developed approach to measure inclination angles with MEMS sensors has been the use of more than one accelerometer to measure the same joint angle [1], [2], [3]. Kalman filtering has also been applied when fusing accelerometer and gyroscope measurements [9], [10]. The developed MEMS sensor prototype, on the other hand, uses complementary filtering [8], [11] to produce angular values by fusing the advantageous static performance of the accelerometer with the favourable dynamic performance of the gyroscope. The accelerometer inclination angle and gyro rate output are governed by

$$y_a = z + \eta_a \quad (1)$$

$$y_g = dz/dt + \eta_g + b \quad (2)$$

where z is the actual inclination angle, η_a is the accelerometer output noise, η_g is the gyro output noise which is predominantly of high frequency like the accelerometer noise term but typically has less variance, and b is a bias term which is predominantly low frequency disturbance.

The complementary filter is implemented with a linear feedback system as shown in Fig. 3 where the combined output of the accelerometer inclination angle y_a and the gyro angular velocity output y_g has the Laplace representation

$$z_{est} = [C(s)y_a]/[s + C(s)] + [y_g s]/[s^2 + C(s)s] \quad (3)$$

where C(s) can be chosen as the PI-controller, i.e. $C(s) = K_P + K_I/s$ so that $G(s) = C(s)/[s+C(s)]$ is a low-pass filter, and its complement $1 - G(s) = s/[s+C(s)]$ is a high-pass filter. In the

PI-controller case, the time domain dynamics are governed by

$$dz_{est}/dt = y_g - b_i + K_P(y_a - z_{est}) \quad (4)$$

$$db_i/dt = -K_I(y_a - z_{est}) \quad (5)$$

where K_P defines the crossover frequency chosen as a trade-off based on the low-pass characteristics of the accelerometer-based inclination and the low frequency bias characteristics of the gyroscope rate measurements. A non-zero K_I rejects a load disturbance b_i from the output.

The time domain operation of the complementary filter is illustrated in Fig. 4, where the initial dominance of the gyro followed by its slowly decaying effect, and conversely, the minor initial effect of the accelerometer followed by steady-state dominance show clearly. The resultant output has a typical steady-state error within ±0.05 degrees whereas the dynamic error is directly proportional to movement velocity. The high accuracy reference sensor output shown in solid grey color was from a Heidenhain incremental encoder with a resolution of 0.00075 degrees per pulse.

III. Condition Monitoring with Unscented Kalman Filter Using MEMS Sensor

Model-based fault detection and isolation (FDI) with UKF was shown in [7], [12] by the authors in a proportional valve controlled 1-DOF hydraulic application in which leakages were emulated and valve faults were simulated under varying load. Now we consider using the developed robust, low-cost MEMS sensor in place of the high accuracy reference sensor in the same application for measuring joint angle and angular velocity. In this section, the procedure for applying the UKF to FDI is described, but it is to be noted that the paper focuses mainly on comparing fault detection capabilities of the different sensors by residual analysis.

A. Principle of unscented Kalman filter-based condition monitoring

Condition monitoring with the UKF is an observer-based method belonging to the class of model-based methods. The model-based condition monitoring is established around the concept of analytical redundancy which originates from running the plant model in parallel to the real physical system, see Fig. 5. The scheme is to subtract plant model output $\hat{y}(k)$ (in this case an UKF estimate) from actual measurement $y(k)$ to create a residual $r(k)$ that reveals the system condition [13]. The residual should be close to zero in faultless operation, assuming that measurement noise is minor and the plant model is accurate. After the appearance of a fault, if the residual is sensitive to the fault, the residual differs from zero thus enabling fault detection. Following fault detection, the cause of the fault is isolated by examining the residual for known patterns. For example in hydraulic systems, external leakages are known to mainly affect the pressure residual of cylinder chamber A or B depending on whether the leakage is on the A or B side, whereas internal leakages affect both the A and B pressure residuals.

The benefits of using UKF in the FDI of hydraulic systems are increased robustness towards process and measurement noise, and the suitability for nonlinear systems if compared with linear FDI filters. It is suitable for general problems with nonlinear process and measurement models

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \quad (6)$$

$$\mathbf{y}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{v}_{k+1} \quad (7)$$

where $\boldsymbol{x}_{k+1}$ is a state vector, $\boldsymbol{f}(\ldots)$ is the process model, $\boldsymbol{u}_k$ is a control vector, $\boldsymbol{w}_k$ is a process noise vector, $\boldsymbol{y}_{k+1}$ is a measurement vector, $\boldsymbol{h}(\ldots)$ is the measurement model, $\boldsymbol{v}_{k+1}$ is a measurement noise vector, and k is a discrete time step.

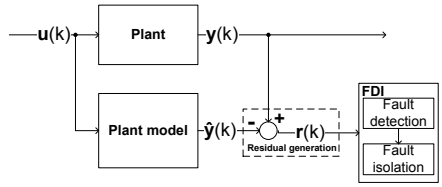


Figure 5. Principle of model-based condition monitoring [13].

The increased robustness of the UKF leads to a non-drifting state estimate in case of bounded short term modeling errors or measurement noise. The nonlinearity property of the UKF allows the modeling of the inherent nonlinearities of hydraulic systems, such as actuator friction, and valve properties including turbulent flow, dead zone, hysteresis, saturation and nonlinear opening. When estimating the states of the above system, the UKF is also typically more efficient than for example a particle filter, which requires the re-sampling of numerous particles at the cost of computational complexity.

To gain efficiency, the UKF carries out an unscented transformation (UT), a technique using deterministic sampling for estimating the mean and covariance of states after they have been transformed through nonlinear equations. The UKF thus avoids the linearization which is partly responsible for estimation errors for example in many extended Kalman filter (EKF) applications. The use of UKF over EKF is further justified by its usually increased accuracy in estimating the true mean and covariance of states [14], [15]. The UKF algorithm is summarized in the Appendix.

B. Reduced-order model of a proportional valve controlled hydraulic cylinder

The reduced state vector of the hydraulic system consisting of a fast proportional valve controlling a hydraulic cylinder is

$$\boldsymbol{x} = [p_A, p_B, x_s, dx_s/dt]^T \quad (8)$$

where p_A is the pressure in cylinder chamber A (piston-side), p_B is the pressure in cylinder chamber B (rod-side), x_s is the spool position of the valve, and dx_s/dt is the spool velocity.

The control vector of the system is:

$$\boldsymbol{u} = [x, dx/dt, u_v, p_s]^T \quad (9)$$

where x and dx/dt are the piston position and velocity, u_v is valve control signal, and p_s is the supply pressure. Note that the piston position and velocity can be obtained via geometrical transformations from the joint angle and angular velocity measurements.

The continuous-time dynamics for the system are

$$\begin{aligned}
\dot{x}_1 &= f_1(x_k,u_k) = \frac{B_{effA}}{V_{0A}+A_A u_1}\left(Q_A - A_A u_2\right)\\
\dot{x}_2 &= f_2(x_k,u_k) = \frac{B_{effB}}{V_{0B}+A_B(x_{\max}-u_1)}\left(Q_B + A_B u_2\right)\\
\dot{x}_3 &= f_3(x_k,u_k) = x_4\\
\dot{x}_4 &= f_4(x_k,u_k) = K\omega_n u_3 - 2\omega_n d_r x_4 - \omega_n^2 x_3
\end{aligned} \tag{10}$$

where B_{effX} is the effective bulk modulus in chamber X, V_{0X} is the chamber volume X, A_X is the cross-sectional cylinder area in chamber X, Q_X is the flow to and from chamber X, for $X = \{A, B\}$, x_{max} is the cylinder stroke, K is the gain from valve control signal to spool position, ω_n is the natural frequency of the spool, and d_r is the damping ratio. Note that the piston position and velocity are treated as inputs as described by (9). Thus the piston dynamics can be excluded from the system model described by (10), which guarantees independence from load force and mass.

The flow rates Q_A and Q_B from (10) are as follows

$$Q_A = \begin{cases}
K_{vPA}x_s\sqrt{|p_s - p_A|}\,\mathrm{sgn}(p_s - p_A), x_s \geq 0.01\\
K_{vAT}x_s\sqrt{|p_A - p_T|}\,\mathrm{sgn}(p_s - p_A), x_s \leq -0.01\\
K_{vPAleak}(p_s - p_A)\\
-K_{vATleak}(p_A - p_T), |x_s| < 0.01
\end{cases} \tag{11}$$

$$Q_B = \begin{cases}
-K_{vBT}x_s\sqrt{|p_B - p_T|}\,\mathrm{sgn}(p_B - p_T),\\
x_s \geq 0.01\\
-K_{vPB}x_s\sqrt{|p_S - p_B|}\,\mathrm{sgn}(p_s - p_B),\\
x_s \leq -0.01\\
K_{vPBleak}(p_s - p_B)\\
-K_{vBTleak}(p_B - p_T), |x_s| < 0.01
\end{cases} \tag{12}$$

where p_T is the tank pressure, K_{vX} and $K_{vX,leak}$ are the flow coefficient and leakage flow coefficient, respectively, of notch X, for $X = \{PA, AT, BT, PB\}$.

By discretizing (10) with Euler's forward method, the states are transferred a time step forward with

$$x_{n,k+1} = x_{n,k} + f_n(\boldsymbol{x}_k, \boldsymbol{u}_k)\Delta T \tag{13}$$

where n = 1,2,...,4, k is the time index, and ΔT is the fixed time step. As explained in [7], there was a linear relation between states and measurements:

$$\hat{\mathbf{y}}_{k+1}^{-} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \hat{\mathbf{x}}_{k+1}^{-} \tag{14}$$

IV. Experiments

A. Test bed and measurements

Using the model presented in Section III.B, the UKF is designed for a 1-DOF hydraulic test bed shown in Fig. 6. The load masses of the system are set to 494 kg on the left, zero on the right, and additionally, the boom weighs 297 kg.

Shown in the hydraulic diagram in Fig. 7, the proportional 4/3-directional valve (24 L/min @ 3.5 MPa) controls the flow to the asymmetrical cylinder (Ø80/45-545). The restrictor valves (labeled 'Internal leakage', 'External leakage A' and 'External leakage B') are needle valves (30 L/min @ 40 MPa) which emulate leakage faults that often in reality require halting system operation after FDI to minimize damage to the environment. The cylinder chamber and the supply pressures are measured with pressure transmitters having a 0-25 MPa operating range. The control signal of the valve is also measured, as are the boom joint angle θ and boom joint angular velocity dθ/dt that are measured with the reference sensor (Heidenhain encoder), and the discussed MEMS sensor prototype that is installed near the rotating joint of the test bed to avoid any perturbations in the accelerometer readings due to centrifugal forces.

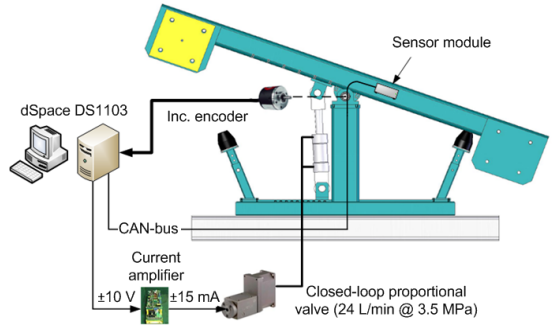


Figure 6. Illustration of test bed setup.

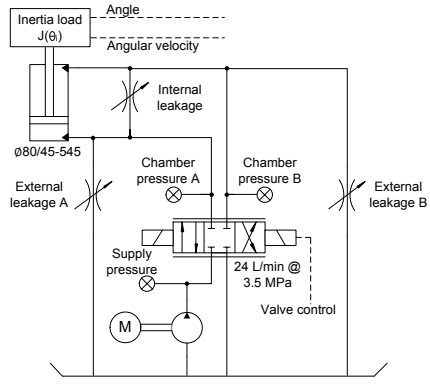


Figure 7. Hydraulic diagram of test bed.

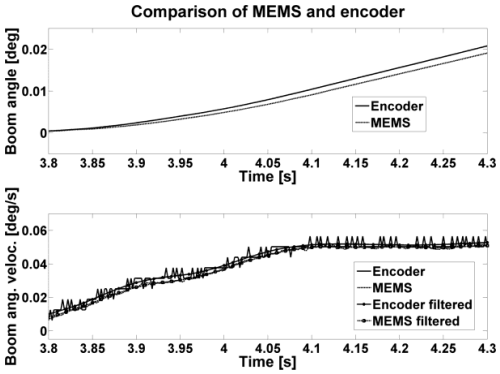


Figure 8. Angle and angular velocity measurements from the angle encoder (solid), MEMS sensor (dashed), filtered encoder angular velocity (asterisk), and filtered MEMS sensor angular velocity (square).

A comparison of the encoder and MEMS measurements in a 0.50-second-window (Fig. 8) shows the performance difference between the sensors. One can see a steady-state error in the MEMS angle measurement and noise in the differentiated angular velocity of the encoder. The angle and angular velocity measurements from both sensors were converted to piston position and velocity and were then used as inputs to the UKF.

B. *Experimental comparison of MEMS and high accuracy encoder in condition monitoring*

A random user-input to the proportional valve was used in all fault-free and fault experiments while the boom angle was measured with both the MEMS sensor and encoder to facilitate a direct comparison between the sensor performances in this condition monitoring task. The performances of the MEMS sensor and angle encoder were compared in terms of pressure residual behavior by using the both sensor outputs as control inputs to the UKF. The residuals were calculated with

$$r_k = p_k - p_{est,k} \tag{15}$$

where p_k is the measured and $p_{est,k}$ the UKF estimated cylinder chamber pressure at time instant k. Pressure residuals were chosen over velocity and position residuals on the basis of their high sensitivity to leakages in the so-called global sensitivity analysis and to enable the important property of load independence; see [12]. The residuals were averaged within a moving window of 5 seconds, and they were paired with adaptive, pressure-dependent thresholds that were required to be crossed to induce an alarm to increase robustness against false alarms; see also [7]. The positive thresholds for the encoder and MEMS sensor systems, respectively, were experimentally set as follows

$$t_{pos,E}(\mu_{p,k}) = 0.015 + 4*10^{-4}\mu^2_{p,k} \tag{16}$$

$$t_{pos,M}(\mu_{p,k}) = 0.023 + 4*10^{-4}\mu^2_{p,k} \tag{17}$$

where μ_p is the cylinder chamber pressure averaged over a 5-second window. On the basis of experimental verification and fault-free experiments conducted in the next section, the first term of the MEMS threshold was set 53 percent higher than the first term of the encoder threshold to minimize false alarms. The process noise variances of the UKF in the MEMS system could have also been increased but that was not considered to allow a fair comparison between the sensors. The negative thresholds were derived from (15) and (16) by simple multiplication with minus one. Note that while this monitoring strategy based on moving average is suboptimum from the viewpoint of minimizing the delay for detection for a fixed mean time between false alarms, see e.g. [16], it allows simple comparison of the different sensors.

C. *Fault-free experiments*

In fault-free experiments, the residuals should stay within the thresholds, and preferably near zero to enable reliable condition monitoring. The premise is that the residual of the MEMS sensor is larger than that of the encoder due to the limited accuracy of the MEMS sensor. To verify this, a total of six fault-free experiments were used, one of which is shown in Fig. 9. From the figure, it can be seen that absolute values of the MEMS residuals are larger than encoder residuals, which lead to the higher threshold requirement for the MEMS sensor system discussed previously. For a numerical comparison between the different sensors, Table I shows the mean of the RMS' of the residuals calculated over the six fault-free experiments of varying durations from approximately 20 to 70 seconds and weighed in proportion to these durations. The results show that on average the MEMS pressure A residuals were 6 % and the pressure B residuals 59 % higher than encoder residuals. This implies poorer fault detection capability for the MEMS sensor system compared to the more accurate encoder system.

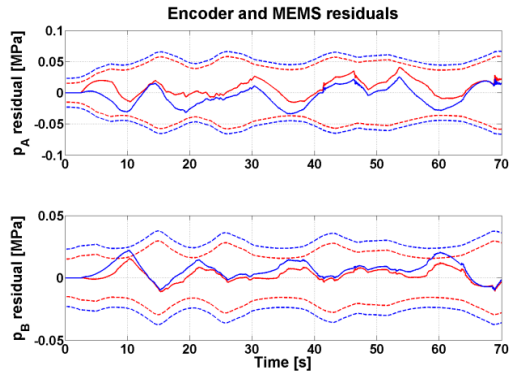


Figure 9. Pressure residuals with the encoder (red) and the MEMS sensor (blue) measurements in a fault-free experiment, the encoder thresholds in dashed red, MEMS thresholds in dashed blue.

TABLE I. RMS' CALCULATED FROM MEMS AND ENCODER MEASUREMENTS IN FAULT-FREE EXPERIMENTS

Sensor	Residual	RMS [MPa]
Encoder	p_A	0.0141
	p_B	0.0051
MEMS	p_A	0.0150
	p_B	0.0081

D. Fault case 1: External leakage in cylinder chamber B

Differences in fault detection capability between the encoder and MEMS sensor systems were studied with external leakage fault cases. An external leakage in cylinder chamber B (rod-side) was added to the system at the 26th second, shown in Fig. 10 with the solid vertical line, by opening the corresponding needle valve. After the introduction, the leakage magnitude was time-variant, see Section IV.F. The fault detection took over ten seconds with both the MEMS sensor and encoder since the residuals were forced to zero when the cylinder was completely retracted due to significant estimation errors in this situation. Overall, the MEMS sensor system was 6 seconds slower than the encoder system in fault detection, and at the time of detection the external leakage was approximately 0.2 L/min and 1.05 L/min with the encoder and the MEMS, respectively. The difference seems remarkable, but in reality the 0.2 L/min was 5.5 % of the flow passing through the proportional valve and the 1.05 L/min was 11 %.

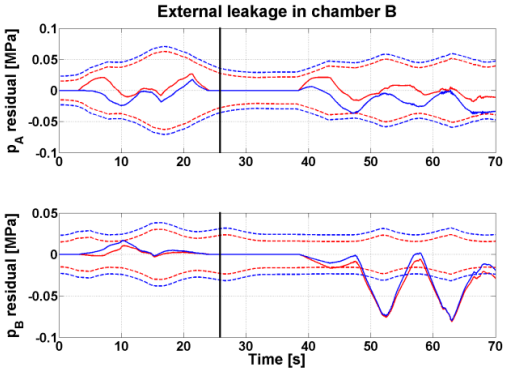


Figure 10. Pressure residuals with the encoder (red) and the MEMS sensor (blue) measurements when there is an external leakage in chamber B (rod-side), the encoder thresholds in dashed red, MEMS thresholds in dashed blue.

E. Fault case 2: Internal leakage in cylinder

Differences in fault detection capability between the encoder and MEMS sensor systems were also studied with internal leakage fault cases. An emulated internal leakage between cylinder chambers that was added to the system around the 30th second is shown in Fig. 11. The fault was detected approximately only 0.3 seconds sooner with the encoder than with the MEMS sensor. At the time of detection the leakage was 1.6 L/min and 1.5 L/min with the encoder and MEMS sensor, respectively. Of the flow through the proportional valve, the leakages were 21 % with the encoder and 19.5 % with the MEMS. Although the MEMS sensor system recognized a lower leakage, it performed worse by failing to recognize the previous 1.6 L/min leakage that was recognized successfully by the encoder system. The recognition of the lower leakage by the MEMS system was due to the delayed fault detection caused by the higher thresholds of the MEMS system, and the time-varying leakage to be discussed in Section IV.F.

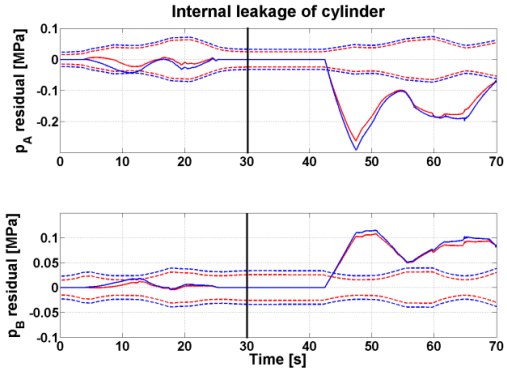


Figure 11. Pressure residuals with the encoder (red) and the MEMS sensor (blue) measurements when there is an internal leakage between cylinder chambers, the encoder thresholds in dashed red, MEMS thresholds in dashed blue.

F. Summary of fault experiments

In Sections IV.D and IV.E, the fault detection performance of the encoder and MEMS sensor were compared with the fault detection delay and with the proportion of the leakage flow rate to the total flow rate through the proportional valve at the time of detection. The MEMS sensor had generally higher fault detection delays than the encoder mainly due to its mandatory higher threshold selection based on the higher residual RMS' in fault-free experiments. The fault detection delay of the MEMS sensor was also greater in the external leakage experiment than in the internal leakage experiment because the internal leakage was generally of higher amplitude and thus resulted in higher absolute values of residuals with both sensor systems. Namely, the external leakage flow rate was 0.38±0.33 L/min with a maximum value of 1.52 L/min and a minimum of 0.04 L/min, whereas the internal leakage flow rate was 0.91±0.44 L/min with a single maximum of 6.49 L/min due to high-amplitude pressure oscillation and a minimum of 0.05 L/min. The proportions of the leak flow rates to the flow rates through the proportional valve at the fault detection instant were comparable between the two sensor systems.

TABLE II. RMS' CALCULATED FROM MEMS AND ENCODER MEASUREMENTS IN FAULT EXPERIMENTS.

Sensor	Fault	Residual	RMS [MPa]
Encoder	External leakage B	p_A	0.0119
		p_B	0.0388
Encoder	Internal leakage	p_A	0.1392
		p_B	0.0817
MEMS	External leakage B	p_A	0.0223
		p_B	0.0365
MEMS	Internal leakage	p_A	0.1534
		p_B	0.0868

To further compare the different sensors, RMS' of the residuals were calculated in fault experiments in a similar fashion as for Table I. These RMS' in Table II show that the operation of the encoder and MEMS sensor systems are

highly comparable in the internal leakage fault cases, but that the encoder performs better in the external leakage fault case by showing a smaller change in the p_A residual than the MEMS sensor, whereas p_B residual increases similarly with both sensor systems. Overall, the RMS' of the different sensors are reasonably well comparable and if compared with the Table I, they provide discriminability of sufficient degree. That is, the MEMS sensor system is capable of functioning in a condition monitoring system when the leakages are fairly minor, though the lowest detectable leakage was not sought here. If further considering the changes in the observed residuals and their signs in the Figs. 10 and 11 when a threshold is exceeded, monitoring the different combinations would allow one to conduct FDI for specifying the origin of the fault.

V. Discussion

The experimental results in the condition monitoring task of detecting oil leakages reveal the limited accuracy of the MEMS-based sensor when compared with the high-accuracy reference, the encoder, but the differences between the two sensor systems are not as drastic as what might be expected based on for example steady-state characteristics. A comparison of steady-state accuracies namely suggests that the encoder is ideally almost 67 times more accurate than the MEMS-based sensor with their accuracies 0.00075 ° and about 0.05 ° empirically found in [8], respectively. But in the condition monitoring, we found that fault detection was 0.3 to 6 seconds slower with the MEMS when the emulated oil leakages were between 0.2 to 1.6 L/min, RMS' of pressure residuals were 6 to 59 percent higher with MEMS sensor than with the encoder in fault-free experiments, and the RMS' in fault experiments were comparable to those. These figures do not resemble the remarkable difference in the estimated steady-state accuracies, which says that the limited accuracy of the MEMS sensor in the form of minor delay and bias does not cause modeling errors that would render it unsuitable for the condition monitoring task. In general, the MEMS measurement is corrupted by multiple error sources with complex mutual dependencies where, for example, scale factor and misalignment contributions depend on the boom dynamics. As the results show, the practical challenge related to the modeling of such error sources can be successfully overcome by the appropriate PI-type complementary filter, UKF covariance parameters, and the pressure residual threshold selection. Thus the MEMS may be considered as a cost-efficient, accurate enough tool for 1-DOF condition monitoring applications.

VI. Conclusion

In this paper, from the two-fold mechatronic perspective, firstly the accuracy and algorithms of the developed MEMS-based motion sensor prototype were presented and discussed. Secondly, the MEMS sensor was applied to an UKF-based condition monitoring scenario of a joint driven by a hydraulic cylinder. In that scenario, angle and angular velocity measurements provided by the developed MEMS sensor prototype and a high-accuracy reference sensor (joint angle encoder) were fed to the UKF as control inputs. Pressure residuals were then generated using both sensors and compared against each other in terms of behavior in fault-free and external and internal leakage experiments to study their effect on fault detection capability. Overall, the performance of the MEMS sensor was suitable for the condition monitoring task in this open-loop system and cannot be considered to be a more limiting factor than modeling errors. Hence the MEMS sensor is able to replace the expensive reference sensor in the model-based condition monitoring of 1-DOF open loop systems. In a closed-loop system though, where the effect of a fault must be compensated by reconfiguring the controller, the prolonged fault detection might have severe consequences.

Appendix

The equations of the unscented Kalman filter algorithm are [15]:

$$\begin{aligned}\hat{\mathbf{x}}_0 &= \mathrm{E}(\mathbf{x}_0)\\ \mathbf{P}_0 &= \mathrm{E}\left[(\mathbf{x}_0-\hat{\mathbf{x}}_0)(\mathbf{x}_0-\hat{\mathbf{x}}_0)^{\mathrm{T}}\right]\\ k &= 0\end{aligned} \tag{18}$$

$$\begin{aligned}\hat{\mathbf{x}}_k^{(0)} &= \hat{\mathbf{x}}_k\\ \hat{\mathbf{x}}_k^{(i)} &= \hat{\mathbf{x}}_k + \tilde{\mathbf{x}}^{(i)}, i = 1,2,\ldots 2N\\ \tilde{\mathbf{x}}^{(i)} &= \sqrt{(N+\lambda)\mathbf{P}_k}_i^{\mathrm{T}}, i = 1,2,\ldots,N\\ \tilde{\mathbf{x}}^{(i)} &= -\sqrt{(N+\lambda)\mathbf{P}_k}_i^{\mathrm{T}}, i = N+1, N+2,\ldots 2N\end{aligned} \tag{19}$$

$$\hat{\mathbf{x}}_{k+1}^{(i)} = \mathbf{f}(\hat{\mathbf{x}}_k^{(i)}, \mathbf{u}_k) \tag{20}$$

$$\begin{aligned}w_{mean}^{(0)} &= \frac{\lambda}{N+\lambda}\\ w_{mean}^{(i)} &= \frac{1}{2(N+\lambda)}, i = 1,2,\ldots 2N\\ \hat{\mathbf{x}}_{k+1}^{-} &= \sum_{i=0}^{2N} w_{mean}^{(i)} \hat{\mathbf{x}}_{k+1}^{(i)}\end{aligned} \tag{21}$$

$$\begin{aligned}w_{\mathrm{cov}}^{(0)} &= \frac{\lambda}{N+\lambda} + (1-\alpha^2+\beta)\\ w_{\mathrm{cov}}^{(i)} &= \frac{1}{2(N+\lambda)}, i = 1,2,\ldots 2N\\ \mathbf{P}_{k+1}^{-} &= \sum_{i=0}^{i=2N} w_{\mathrm{cov}}^{(i)} (\hat{\mathbf{x}}_{k+1}^{(i)} - \hat{\mathbf{x}}_{k+1}^{-})(\hat{\mathbf{x}}_{k+1}^{(i)} - \hat{\mathbf{x}}_{k+1}^{-})^{\mathrm{T}} + \mathbf{Q}_k\end{aligned} \tag{22}$$

$$\begin{aligned}\hat{\mathbf{y}}_{k+1}^{(i)} &= \mathbf{h}(\hat{\mathbf{x}}_{k+1}^{(i)}, \mathbf{u}_k)\\ \hat{\mathbf{y}}_{k+1} &= \sum_{i=0}^{2N} w_{mean}^{(i)} \hat{\mathbf{y}}_{k+1}^{(i)}\\ \mathbf{P}_{yy} &= \sum_{i=0}^{2N} w_{\mathrm{cov}}^{(i)} (\hat{\mathbf{y}}_{k+1}^{(i)} - \hat{\mathbf{y}}_{k+1})(\hat{\mathbf{y}}_{k+1}^{(i)} - \hat{\mathbf{y}}_{k+1})^{\mathrm{T}} + \mathbf{R}_{k+1}\end{aligned} \tag{23}$$

$$\mathbf{P}_{xy} = \sum_{i=0}^{2N} w_{\text{cov}}^{(i)} (\hat{\mathbf{x}}_{k+1}^{(i)} - \hat{\mathbf{x}}_{k+1})(\hat{\mathbf{y}}_{k+1}^{(i)} - \hat{\mathbf{y}}_{k+1})^{\mathrm{T}} + \mathbf{R}_{k+1} \quad (24)$$

$$\mathbf{K}_{k+1} = \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1} \quad (25)$$

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1}^{-} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1}) \quad (26)$$

$$\mathbf{P}_{k+1} = \mathbf{P}_{k+1}^{-} - \mathbf{K}_{k+1}\mathbf{P}_{yy}\mathbf{K}_{k+1}^{\mathrm{T}} = \mathbf{P}_{k+1}^{-} - \mathbf{P}_{xy}(\mathbf{P}_{yy}^{-1})^{\mathrm{T}}\mathbf{P}_{xy}^{\mathrm{T}} \quad (27)$$

where E is the expectation operator, $\boldsymbol{x}_k$ is the true state vector, $\hat{\mathbf{x}}_k^-$ is the model prediction of the state vector, $\hat{\mathbf{x}}_k$ is the state estimate vector, k is a discrete time instant variable, $\tilde{\mathbf{x}}^{(i)}$ is the sigma point vector, $\boldsymbol{P}_k$ is the state error covariance matrix, $\boldsymbol{f}$ is a nonlinear function for transferring states to the next time instant, $\boldsymbol{h}$ is a nonlinear function for converting states to measurements, w is a weighting coefficient, N is the dimension of the state vector and λ is a scaling parameter, satisfying $\lambda=\alpha^2 (L+\kappa) - N$. The parameter α, with a typical value of 10^{-3}, is a tuning factor determining the spread of the sigma points. The constant κ is a secondary tuning parameter, usually zero. The constant β affects the weight of the first error covariance term, and the choice $\beta = 2$ is optimal for normally distributed states. Cholesky decomposition should be used for the matrix square root in (16) for computational efficiency.

An iteration of the UKF algorithm proceeds as follows:

1. Initialize, (18)
2. Estimate the a priori state vector $\hat{\mathbf{x}}_{k+1}^-$ (prediction)
   a. Generate sigma points around the previous estimate, (19)
   b. Propagate the sigma points through the nonlinear functions, (20)
   c. Calculate the state mean, (21)
3. Calculate the a priori error covariance $\mathbf{P}_{k+1}^-$, (22)
4. Estimate the a posteriori state vector $\hat{\mathbf{x}}_{k+1}$
   a. Unscented transformation of predicted measurements for capturing the mean and covariance, (23)
   b. Calculate the cross-covariance between predicted states and measurements, (24)
   c. Calculate the Kalman gain, (25)
   d. Update state estimate, (26)
5. Calculate the a posteriori error covariance $\mathbf{P}_{k+1}$, (27)
6. Return to step 2

REFERENCES

[1] F. Ghassemi, S. Tafazoli, P. D. Lawrence, and K. Hashtrudi-Zaad, "An Accelerometer-Based Joint Angle Sensor for Heavy-Duty Manipulators," *Proceedings of the IEEE International Conference on Robotics and Automation*. Washington DC, May 2002, pp. 1771 - 1776.

[2] F. Ghassemi, P. Lawrence, and K. Hashtrudi-Zaad, "Design and calibration of an integration-free accelerometer-based joint-angle sensor," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 1, 2008.

[3] P. Cheng, F. Linnarsson, and B. Oelmann, "Joint Angular Sensor Based on Distributed Biaxial MEMS Accelerometers," *Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON)*, Taipei, Nov. 2007, pp. 2242 - 2247.

[4] J. Leavitt, A. Sideris, and J. E. Bobrow, "High bandwidth tilt measurement using low-cost sensors," *IEEE/ASME Transactions on Mechatronics*, vol. 11, no. 3, 2006.

[5] G. S. Maruthi, and K. Panduranga Vittal, "Electrical fault detection in three phase squirrel cage induction motor by vibration analysis using MEMS accelerometer," *Proc. of the International Conf. on Power Electronics and Drives Systems*, 2005, vol. 2, pp. 838-843.

[6] A. Albarbar, S. Mekid, A. Starr and R. Pietruszkiewicz, "Suitability of MEMS accelerometers for condition monitoring: an experimental study," *Sensors*, vol. 8, pp. 784-799, 2008.

[7] J. Nurmi, and J. Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 2: Fault Detection and Isolation scheme," *Int. Journal of Fluid Power*, vol. 13, no. 1, March 2012, pp. 17-27.

[8] J. Honkakorpi, J. Vihonen and J. Mattila, "Sensor module for hydraulic boom state feedback control", Int. Journal of Fluid Power, vol. 13, no. 3, pp. 15-23, Nov. 2012.

[9] F. J. O. Corpuz, B. C. Y. Lafoteza, R. A. L. Broas, and M. Ramos, "Design and Implementation of a Closed-Loop Static Balance System for the YICAL Leg 2 Biped," *TENCON 2009 IEEE Region 10 Conference*, Jan. 2009, Singapore.

[10] M. Quigley, R. Brewer, S. P. Soundararaj, V. Pradeep, Q. Le, and A. Y. Ng, "Low-cost Accelerometers for Robotic Manipulator Perception", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, Oct. 2010.

[11] R. Mahony, T. Hamel, and J-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203-1218.

[12] J. Nurmi, and J. Mattila, "Detection and isolation of leakage and valve faults in hydraulic systems in varying loading conditions, Part 1: Global Sensitivity Analysis," *Int. Journal of Fluid Power*, vol. 12, no. 3, Nov. 2011, pp. 41-51.

[13] R. Isermann, *Fault-diagnosis systems, an introduction from fault detection to fault tolerance*. Springer-Verlag Berlin Heidelberg, 2006, pp. 1- 7.

[14] S. J. Julier, and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," *Proc. of AeroSense, the 11th international symposium on aerospace/defence sensing, simulation and controls*, 1997, pp. 182-193.

[15] E. A. Wan, and R. van der Merwe, "The unscented Kalman filter for nonlinear estimation," *Proc. of IEEE adaptive systems signal processing, communication and control symposium*, October 2000, pp. 153–158.

[16] M. Basseville, and I. V. Nikiforov, *Detection of abrupt changes: theory and application.* Prentice-Hall, 1993, pp. 25-40.

Publication IV

Jarmo Nurmi, Jouni Mattila, "Detection and isolation of faults in mobile hydraulic valves based on a reduced-order model and adaptive thresholds," *Proceedings of the ASME/BATH Symposium on Fluid Power & Motion Control (FPMC)*, Sarasota, Florida, USA, October 6–9, 2013, http://dx.doi.org/10.1115/FPMC2013-4435.

Publication V

Jarmo Nurmi, Jouni Mattila, "Global energy-optimised redundancy resolution in hydraulic manipulators using dynamic programming," *Automation in Construction*, vol. 73, pp. 120 – 134, January 2017, http://dx.doi.org/10.1016/j.autcon.2016.09.006.

Automation in Construction 73 (2017) 120–134

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Global energy-optimised redundancy resolution in hydraulic manipulators using dynamic programming

Jarmo Nurmi*, Jouni Mattila

Tampere University of Technology, Department of Intelligent Hydraulics and Automation, P.O. Box 589,Korkeakoulunkatu 6, Tampere FIN-33101, Finland

ARTICLE INFO

Article history:
Received 16 January 2015
Received in revised form 2 September 2016
Accepted 20 September 2016
Available online 30 September 2016

Keywords:
Redundancy resolution
Hydraulic manipulator
Construction crane
Energy optimisation
Global optimisation
Dynamic programming
Joint limits
Load-sensing system
Constant-pressure system

ABSTRACT

This paper addresses the problem of redundancy resolution in closed-loop controlled hydraulic manipulators. The problem is treated at the hydraulic level using proposed cost functions formulated into a dynamic programming approach of minimum-state representation. Bounds on joint range, actuator velocity and acceleration were enforced. This approach minimises the hydraulic energy consumption of the widely popular load-sensing and constant-supply pressure systems. The presented approach can resolve the redundancy more effectively from the hydraulic side than do actuator velocity or energy optimisation approaches, point-wise optimal approaches or some standard direct optimisation tools that may lead to inferior solutions, as shown in simulation results where up to 15–30% greater energy use is seen with some competing approaches. The results obtained motivate joint trajectory optimisation at the hydraulic level in prospective applications at construction sites where frequently driven work cycles of hydraulic construction cranes are automated.

1. Introduction

Hydraulic manipulators are widely used for excavation and lifting applications at construction sites and for heavy-duty material handling in the forest industry due to their superior power-density and rugged nature. Although the hydraulic construction cranes are mainly open-loop controlled by human operators, manufacturers in these industries are interested in broadening their offerings through the automation of typical work cycles to improve the productivity and safety of their machines. On a technical level, this automation requires solving the inverse kinematics problem and realising closed-loop control (see [1] for the proposed closed-loop control algorithm). Because construction cranes are typically equipped with redundant joints, the inverse kinematics problem transforms into a more difficult redundancy resolution problem, thus lending itself to sophisticated machine operation optimisation. Here, we resolve the redundancy of the construction crane from the standpoint of hydraulic energy minimisation. This approach of redundancy resolution entails moving the crane cylinder actuators in an energy-efficient fashion that is also subject to task-space reference.

Only a handful of articles discuss the redundancy resolution of hydraulic manipulators, including [2], in which point-wise optimal joint trajectories are given that minimise the energy consumed by hydraulic actuators. This point-wise solution is sub-optimal over the entire trajectory, and the problem is not fully considered at the hydraulic system level. In [3], some productivity problems in hydraulic knuckle booms are solved locally using redundancy to maximise the lifting capacity or velocity. Dynamic programming is also used to globally minimise the time required to move between two points in the workspace. However, no energy-related objectives were discussed. In [4], the working cycle duration of non-redundant excavators is reduced locally by maximising its joint velocities. Although the article was written from a hydraulics standpoint, energy optimisation was disregarded.

In contrast, much work has been dedicated to resolving the redundancy of general manipulators (e.g. [5,6]). Many of these papers discuss resolved redundancy pertaining to the minimisation of actuator energy consumption, which can lead to significant energy savings in manipulators in general. However, this solution is inevitably sub-optimal when dealing with many hydraulic system types. This generally arises from the pressure losses encountered in

* Corresponding author.
E-mail addresses: jarmo.nurmi@tut.fi (J. Nurmi), jouni.mattila@tut.fi (J. Mattila).

http://dx.doi.org/10.1016/j.autcon.2016.09.006
0926-5805/

most hydraulic systems when the actuators are subject to unequal loads. Therefore, the energy optimisation of hydraulic manipulators that are mainly powered by load-sensing or constant-pressure systems calls for effective control approaches specifically tailored for these hydraulic systems and cost functions formulated at the hydraulic level, instead of the actuator level. Furthermore, contemporary articles on redundancy resolution mostly exemplify highly redundant manipulators that do not represent typical hydraulic manipulators, which have less kinematic redundancy. Therefore, the energy savings presented do not equal the savings typically achieved with hydraulic manipulators. Although the problem is simpler than most in terms of redundancy, the nonlinearities and non-convexity make the problem difficult to solve at the hydraulic level. For example, conventional direct optimisation methods yield local optimums, and the search for a global optimum among the local optimums is seen as time consuming.

In this paper, we effectively explore the redundancy resolution problem using popular hydraulic systems powered by constant-supply pressure or load-sensing variable displacement pumps as opposed to ineffective sub-optimal approaches. We focus on a common 3-degree-of-freedom (DOF) hydraulic manipulator design, which is redundant in one DOF in the typical manner, and propose cost functions at the hydraulic level to globally to minimise the manipulator's hydraulic energy consumption over prescribed workspace movements. To effectively resolve the redundancy, the proposed cost functions are formulated into a minimum-state dynamic programming approach, which ensures accurate tracking of a Cartesian path while minimising the said cost functions. Bounds on joint ranges based on cylinder stroke, cylinder velocities and cylinder acceleration are enforced. We investigate popular load-sensing systems and analyse pump flow rate minimisation, which equally minimises the energy consumption of a constant-supply pressure hydraulic system. We compare our results to well-known sub-optimal control strategies. To the authors' knowledge, this is the first time joint trajectories have been globally optimised at the hydraulic level in prescribed Cartesian motions in relation to typical redundant hydraulic manipulators.

This paper is organised as follows. In Section 2, we introduce a typical hydraulic manipulator with a redundant degree-of-freedom and define its end-effector position and velocity. We also discuss the use of variable displacement pumps in the conventional constant-supply pressure and load-sensing systems. In Section 3, we define the optimal control problems in the continuous and discrete form, and we introduce the dynamic programming approach in Section 4. In Section 5, we propose the cost functions at the hydraulic level. In Section 6, we provide numerical simulations to compare and estimate the energy conservation attainable with a typical manipulator. In Section 7, we discuss important aspects of the optimal control problem, and we provide conclusions in Section 8.

2. Hydraulic manipulator with kinematic redundancy

Let us consider the planar 3-DOF hydraulic manipulator shown in Fig. 1, which represents the typical hydraulic manipulator configuration used in a number of applications for tasks involving heavy lifting at construction sites. The manipulator has a prismatic reach actuator that provides an additional DOF. Because of this redundancy property, the manipulator's end-effector tip can be controlled in an infinite number of joint trajectories, from an initial Cartesian point to the desired end point. This desirable redundancy opens up the possibility of finding joint trajectories that globally optimise the energy consumption of the manipulator at the hydraulic level while the end-effector satisfies Cartesian reference path constraints. To this end, the end-effector position and velocity are defined, and

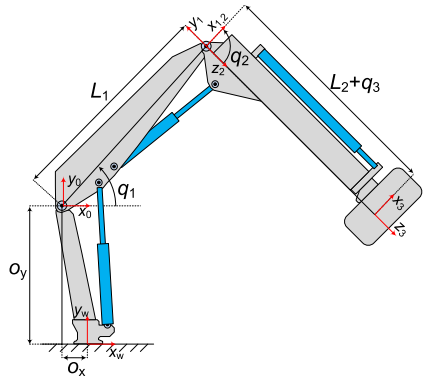


Fig. 1. Typical 3-DOF kinematically redundant hydraulic manipulator.

we discuss the variable displacement pumps heavily utilised in the manipulator's hydraulic systems.

2.1. End-effector position and velocity

The joint space vector of the manipulator is written as

$$\mathbf{q} = [\, q_1 \;\; q_2 \;\; q_3 \,]^{\mathrm{T}} \tag{1}$$

where joint coordinate q_1 denotes the lift angle, joint coordinate q_2 denotes the tilt angle (transfer angle) and the redundant joint coordinate q_3 denotes the extension length of the cylinder (reach). The joint coordinates q_1 (real positive), q_2 (real negative) and q_3 (real positive) are chosen based on the classical Denavit-Hartenberg (DH) convention [7]. Coordinate frames are attached to the links and numbered based on this DH convention. The world coordinate frame in the base is denoted with w (see Fig. 1).

The DH homogeneous transformation matrix $\mathbf{A}_i^{i-1} \in \mathbb{R}^{4\times4}$, which determines the coordinate transformation from the link attached frame i to frame $i-1$, is

$$\mathbf{A}_i^{i-1} = \begin{bmatrix} \mathrm{c}_{\theta_i} & -\mathrm{s}_{\theta_i}\mathrm{c}_{\alpha_i} & \mathrm{s}_{\theta_i}\mathrm{s}_{\alpha_i} & a_i\mathrm{c}_{\theta_i} \\ \mathrm{s}_{\theta_i} & \mathrm{c}_{\theta_i}\mathrm{c}_{\alpha_i} & -\mathrm{c}_{\theta_i}\mathrm{s}_{\alpha_i} & a_i\mathrm{s}_{\theta_i} \\ 0 & \mathrm{s}_{\alpha_i} & \mathrm{c}_{\alpha_i} & \mathrm{d}_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where e.g. s_{θ_i} denotes $\sin(\theta_i)$, c_{θ_i} denotes $\cos(\theta_i)$ and the matrix elements are obtained using the DH parameters (see Table 1). Using the DH transformation matrix in succession, we get

$$\mathbf{A}_3^0 = \mathbf{A}_1^0\mathbf{A}_2^1\mathbf{A}_3^2 \tag{3}$$

where $\mathbf{A}_3^0$ is the total coordinate transformation from the end-effector frame 3 to frame 0. To transform to world frame w,

Table 1
Denavit-Hartenberg parameters of the manipulator.

Joint i	a_i	α_i	d_i	θ_i
1	L_1	0	0	q_1
2	0	$\pi/2$	0	$\pi/2+q_2$
3	0	0	L_2+q_3	0

which is located in the manipulator base, we use a homogeneous transformation matrix on $\mathbf{A}_3^0$

$$\mathbf{T}_3^{\mathrm{w}} = \begin{bmatrix} \mathbf{R}_0^{\mathrm{w}} & \mathbf{o}_0^{\mathrm{w}} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{A}_3^0 \tag{4}$$

where the rotation matrix $\mathbf{R}_0^{\mathrm{w}}$ is the identity matrix $\mathbf{I}_{3\times3}$ because the orientations of coordinate frames 0 and w are aligned, the translation vector $\mathbf{o}_0^{\mathrm{w}}$ is $[o_{\mathrm{x}}\, o_{\mathrm{y}} 0]^{\mathrm{T}}$, with its components denoting the offsets between the frames w and 0, and $\mathbf{T}_3^{\mathrm{w}}$ yields the transformation from frame 3 to w. Finally, we may determine the end-effector position in the world coordinate frame from $\mathbf{T}_3^{\mathrm{w}}$ to be

$$x_{\mathrm{w}} = L_1 \cos(q_1) + (L_2 + q_3)\cos(q_1 + q_2) + o_{\mathrm{x}} \tag{5}$$

$$y_{\mathrm{w}} = L_1 \sin(q_1) + (L_2 + q_3)\sin(q_1 + q_2) + o_{\mathrm{y}} \tag{6}$$

where x_{w} denotes the position of the end-effector in the direction of the x-axis in the world coordinate frame, y_{w} denotes the position of the end-effector in the direction of the y-axis in the world coordinate frame and because the manipulator operates in the xy-plane, the position of the end-effector in the z_{w} coordinate is zero. The manipulator link lengths are L_1 and L_2, and o_{x} and o_{y} are the offset of the world coordinate system in the direction of the x- and y-axes, respectively, from the first joint. The joints are actuated by hydraulic cylinders. The joints q_1 and q_2 are rotational joints, which rotate around the z-axis. The extension joint q_3 is a prismatic joint whose length is determined by the length of the hydraulic cylinder. Because of the physical limits of the hydraulic cylinders that actuate the manipulator joints, the joint coordinates are lower and upper bounded. The maximum joint velocities in both directions of motion are also bounded as dictated by the maximum displacement of the hydraulic pump and cylinder size. System restrictions also limit maximum joint accelerations.

The task-space vector denoting the end-effector Cartesian position may be written as

$$\mathbf{x}_{\mathrm{t}} = \begin{bmatrix} x_{\mathrm{w}} & y_{\mathrm{w}} \end{bmatrix}^{\mathrm{T}} \tag{7}$$

Differentiating $\mathbf{x}_{\mathrm{t}}$ in view of Eqs. (5)–(7) with respect to time yields

$$\dot{\mathbf{x}}_{\mathrm{t}} = \mathbf{J}(\mathbf{q})\dot{\mathbf{q}} \tag{8}$$

where the Jacobian matrix $\mathbf{J}(\mathbf{q}) \in \mathbb{R}^{2\times3}$ is the partial derivative of $\mathbf{x}_{\mathrm{t}}$ with respect to $\mathbf{q}$ and $\dot{\mathbf{q}}$ is the time derivative of $\mathbf{q}$.

To solve the joint velocities $\dot{\mathbf{q}}$ from Eq. (8) we need to invert the non-square matrix $\mathbf{J}(\mathbf{q})$. A particular inversion is obtained using the weighted right-side pseudo-inverse [6]

$$\dot{\mathbf{q}}^{\dagger} = \mathbf{J}(\mathbf{q})^{\dagger}\dot{\mathbf{x}}_{\mathrm{t}} = \mathbf{W}^{-1}\mathbf{J}(\mathbf{q})^{\mathrm{T}}\left(\mathbf{J}(\mathbf{q})\mathbf{W}^{-1}\mathbf{J}(\mathbf{q})^{\mathrm{T}}\right)^{-1}\dot{\mathbf{x}}_{\mathrm{t}} \tag{9}$$

where $\mathbf{W}$ is a weighting matrix and $\mathbf{J}(\mathbf{q})^{\dagger}$ denotes the pseudo-inverse of the Jacobian matrix. This pseudo-inverse minimises the instantaneous Euclidean norm of joint velocities $\dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{q}}$ when the weighting matrix $\mathbf{W}$ equals the identity matrix $\mathbf{I}$, thus yielding a least-square solution. Variations of this weighting approach that are more sophisticated have been presented (e.g. weighting with the manipulator inertia matrix minimises the instantaneous kinetic energy). The pseudo-inverses have also been derived in actuator coordinates [8]. Because these pseudo-inverse approaches lead to the instantaneous minimisation of the performance criteria, the minimum over the whole task-space trajectory must be found using an optimal control methodology.

2.2. Control of hydraulic systems by variable displacement pumps

Variable displacement pumps are hydraulic components capable of outputting a variable flow rate through the hydraulic or electric alteration of the pump displacement. The pumps predominantly used in hydraulic manipulators are axial-piston types in which the output flow is controlled by adjusting the angle of the swash-plate, which is a tilted disc usually actuated by pressure-regulated hydraulic control pistons acting against a spring load. The dynamic characteristics of common hydraulic piston pumps and their control principles are widely known and have been thoroughly studied in the literature (see [9–11]).

These variable displacement pumps are used in constant-supply pressure (CP) and load-sensing systems (LS) in hydraulic manipulators. The pressure and flow are usually regulated through hydro-mechanical feedback realised with a regulator valve subject to hydraulic pressure and spring forces, which control the valve's position. In the CP system, the regulator valve controls the pump displacement in such a way as to maintain a constant supply pressure, regardless of changes in the flow demand or load pressure. The pump flow is simultaneously matched to the required flow of the actuators. This desired control action is accomplished by fixing the spring tension force acting on the spool of the regulator valve to a force corresponding to the desired supply pressure. The hydraulic circuit of the CP system is shown in Fig. 2. CP systems have been widely adopted in hydraulic manipulators used in feedback control applications because of their enhanced stability and simplicity over other systems; however, these benefits come at the cost of energy efficiency.

The more complex LS system shown in Fig. 3 controls the pump displacement through the regulator valve so the supply pressure at the pump outlet tracks a time-varying reference, which is continuously adjusted to a fixed-pressure delta above the highest load pressure sensed from the hydraulic pilot line. Like in the CP system, the pump flow is matched to the actuator requirement. Abiding by this principle, the pressure losses over the valves controlling the actuators may be notably reduced compared to the CP system, but the potential disadvantages are poor energy efficiency with unequal loading and the increased chance of stability problems

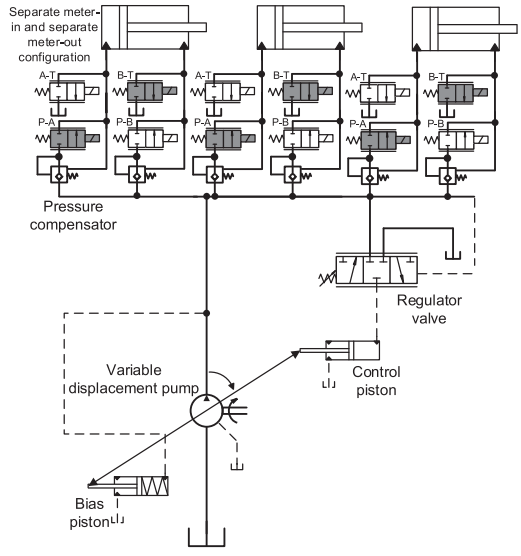


Fig. 2. Hydraulic circuit of a constant pressure system.

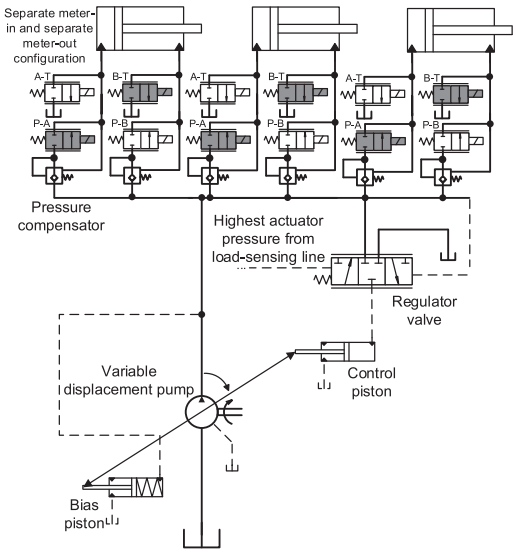


Fig. 3. Hydraulic circuit of a load-sensing system.

because of the decreased damping [12]. Variations of the LS systems include the electrical LS systems [13], which eliminate the long LS pilot line; and flow control systems [14], which remove the pressure feedback on the load at the pump. LS systems are immensely popular in open-loop-controlled manipulators in which energy must be distributed to multiple actuators with a single pumping unit; however, they can also be encouragingly stable in feedback control applications [15].

The vast popularity of these systems in applications and their different operating principles imply the need for a tailored solution to effectively resolve kinematic redundancy at the hydraulic system level. The controllability of the hydraulic pump's flow rate and supply pressure enables this energy saving redundancy resolution.

3. Problem formulation

Let us formulate a dual-objective problem in which the secondary objective is to minimise a performance cost function L_p related to the energy consumption of the hydraulic system (cost functions presented later) while the manipulator end-effector is primarily required to track a time-dependent planar path $\mathbf{r}(t)$ denoted with

$$\mathbf{r}(t) = \left[\, r_\mathrm{x}(t) \;\; r_\mathrm{y}(t) \,\right]^\mathrm{T} \tag{10}$$

where t denotes time and the differentiable Cartesian x and y-coordinate references are $r_\mathrm{x}(t)$ and $r_\mathrm{y}(t)$, respectively.

3.1. Continuous-time formulation

The complex optimal control problem with fixed terminal time is defined as follows:

$$\min_{\mathbf{u}\in\boldsymbol{\rho}} \int_0^{t_\mathrm{f}} L_\mathrm{p}\left(\mathbf{x}(t),\mathbf{u}(t)\right)\, \mathrm{d}t \tag{11}$$

subject to:

$$\begin{aligned}
&\mathbf{x}(0) = \mathbf{x}_0,\\
&\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t),\mathbf{u}(t)),\\
&\mathbf{g}_\mathrm{e}\left(\mathbf{x}(t),\mathbf{u}(t)\right) = 0,\\
&\mathbf{g}_\mathrm{i}\left(\mathbf{x}(t),\mathbf{u}(t)\right) \le 0
\end{aligned} \tag{12}$$

where L_p is the performance cost to be minimised (real), t_f is the terminal time, $\mathbf{x}_0 \in \mathbb{R}^6$ is the initial system state vector, $\mathbf{x} \in \mathbb{R}^6$ denotes the system state vector, $\mathbf{u} \in \mathbb{R}^3$ denotes the control vector, $\mathbf{f} \in \mathbb{R}^6$ denotes the system dynamics, $\mathbf{g}_\mathrm{i} \in \mathbb{R}^{18}$ denotes the inequality constraints, $\mathbf{g}_\mathrm{e} \in \mathbb{R}^2$ denotes the equality constraints and $\boldsymbol{\rho}$ is the control policy search space of feasible $\mathbf{u}(t)$, defined at time indices from 0 to t_f and constrained by joint acceleration limits at each time index (see Eq. (15)).

The system dynamics $\mathbf{f}(\mathbf{x}(t),\mathbf{u}(t))$ are defined using

$$\begin{aligned}
\dot{x}_1 &= x_2\\
\dot{x}_2 &= u_1\\
\dot{x}_3 &= x_4\\
\dot{x}_4 &= u_2\\
\dot{x}_5 &= x_6\\
\dot{x}_6 &= u_3
\end{aligned} \tag{13}$$

where state vector $\mathbf{x} = [q_1 \;\; \dot{q}_1 \;\; q_2 \;\; \dot{q}_2 \;\; q_3 \;\; \dot{q}_3]^\mathrm{T}$ and control vector $\mathbf{u} = [\ddot{q}_1 \;\; \ddot{q}_2 \;\; \ddot{q}_3]^\mathrm{T}$. The time indices were omitted for brevity.

To ensure satisfactory path tracking in the task space, we require that $\mathbf{g}_\mathrm{e}$ contains the time-varying equality constraint

$$\mathbf{x}_\mathrm{t}(t) - \mathbf{r}(t) = 0 \tag{14}$$

However, tracking the path could be equivalently forced with a velocity equality constraint $\dot{\mathbf{x}}_\mathrm{t}(t) - \dot{\mathbf{r}}(t) = 0$, positional inequality constraint $|\mathbf{x}_\mathrm{t}(t) - \mathbf{r}(t)| \le \delta$ with a small constant $\delta \in \mathbb{R}_+$, or, similarly, a velocity inequality constraint. Notice that when using any one of these constraints, we do not need to find weights for the performance and tracking objectives.

The following state and control constrains are contained in $\mathbf{g}_\mathrm{i}$

$$\begin{aligned}
x_{i_\mathrm{min}} &\le x_i \le x_{i_\mathrm{max}}\\
u_{i_\mathrm{min}} &\le u_i \le u_{i_\mathrm{max}}
\end{aligned} \tag{15}$$

where x_{i_min} denotes the minimum feasible value of state $i \in \{1,2,3,4,5,6\}$ (joint position or joint velocity limit), x_{i_max} denotes the maximum feasible value of state $i \in \{1,2,3,4,5,6\}$ (joint position or joint velocity limit), u_{i_min} is the minimum joint acceleration of joint $i \in \{1,2,3\}$ and u_{i_max} is the maximum joint acceleration of joint $i \in \{1,2,3\}$. These inequality constraints are based on physical bounds on joint ranges, joint velocity and joint acceleration. The constraints can be readily reduced into the general form denoted with $\mathbf{g}_\mathrm{i}$ in Eq. (12).

3.2. Discrete-time formulation

The discrete-time problem may be formulated as follows. Firstly, let the continuous time from 0 to t_f be discretised into N intervals of equivalent length t_f/N. Then the discrete version of the cost functional from Eq. (11) may be written as

$$\min_{\mathbf{u}\in\boldsymbol{\rho}} \left\{ T_\mathrm{s} \sum_{k=0}^{N-1} \left[L_{\mathrm{p},k}\left(\mathbf{x}_k,\mathbf{u}_k\right) + \ell_k\left(\mathbf{x}_k,\mathbf{u}_k\right) \right] \right\} \tag{16}$$

where T_s is the integration step, k denotes the discrete time index (stage), $\mathbf{x}_k$ denotes the discrete state vector, $\mathbf{u}_k$ denotes the discrete control vector, $L_{p,k}$ is the discrete performance cost at time stage k, $\ell_k(\mathbf{x}_k, \mathbf{u}_k)$ is the additive term at time stage k that penalises the violation of joint limits (joint ranges, joint velocity and joint acceleration) and $\boldsymbol{\rho}$ is the control policy search space of feasible $\mathbf{u}_k$, defined for time indices from 0 to $N-1$. When the joint limits are violated (based on the limits defined in Eq. (15)), a substantial constant much higher than the normally highest cost function value is added to term ℓ_k to ensure that controls that lead to exceeding joint limits are avoided. The final cost at N, which is independent of the control, is omitted.

The state dynamics may be discretised by using the well-known explicit, forward Euler method

$$\mathbf{x}_{k+1} = \mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k + T_s \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \tag{17}$$

where $\mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k)$ denotes the discretised system dynamics and T_s is the integration step.

The main problems with the continuous and derived discrete formulation are their high-dimensionality and complexity. The high-dimensionality implies that dynamic programming as such is impractical, whereas, for example, the complexity and non-convexity of some of the constraints and cost functions we shall introduce signify that conventional direct optimisation methods would yield locally optimal solutions, depending on the initial guess of the control. The global solution could, however, be searched by iterating through the vast number of solutions generated from different initial guesses; however, this process is ad hoc and time consuming. Nevertheless, because we are searching for a global solution, dynamic programming is a viable candidate, but only with the modified modelling approach suggested in the following section.

4. Dynamic programming solution

Dynamic programming (DP) is a powerful discrete-time method that has one remarkable property in that it can provide a global solution to non-convex optimal control problems. The main disadvantage of DP is its computational complexity: as the number of system states or controls increase, the computational complexity increases exponentially [16]. Hence, dynamic programming can be considered a practical approach only when dealing with low-dimensional problems. DP is based on the well-known principle of optimality defined by Richard Bellman in 1957 [17]:

> An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

In practice, the discretisation of states and controls is required, which yields a grid where the variables involved can take only a finite number of discrete values. The cost-to-go is evaluated only at these discrete points. The finer the discretisation, the closer to the global optimum the solution obtained will be. With the discretisation in place, the solution to our optimal control problem may be obtained using the recursive Bellman equation (principle of optimality):

$$\begin{aligned} J_k(\mathbf{x}_k) = \min_{\mathbf{u}_k \in \mathbb{U}} \{ & L_{p,k}(\mathbf{x}_k, \mathbf{u}_k) + \ell_k(\mathbf{x}_k, \mathbf{u}_k) \\ & + J_{k+1}(\mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k)) \} \end{aligned} \tag{18}$$

where $J_k(\mathbf{x}_k)$ is the optimal cost-to-go from stage $k \in \{0, 1, \ldots, N-1\}$ to the final stage $N-1$ defined for each state vector combination at stage k such that $\mathbf{x}_k \in \mathbb{X} = \{\mathbb{X}_1 \times \mathbb{X}_2 \times \cdots \times \mathbb{X}_{N_x}\}$ (N_x is the number of states) and computed over all control vector combinations at stage k such that $\mathbf{u}_k \in \mathbb{U} = \{\mathbb{U}_1 \times \mathbb{U}_2 \times \cdots \times \mathbb{U}_{N_u}\}$ (N_u is the number of controls). Here, $\mathbb{X}_i = \{x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(N_{x_i})}\}$ and $\mathbb{U}_i = \{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(N_{u_i})}\}$ denote the discrete sets of the state and control i, respectively. The joint limits defined in Eq. (15) determine the maximum and minimum values in these discrete sets of feasible state and control values. The variables N_{x_i} and N_{u_i}, denote the number of discretised states and controls, respectively. The sum $L_{p,k} + \ell_k$ represents the running cost and $\mathbf{F}_k(\mathbf{x}_k, \mathbf{u}_k)$ is the discretised version of system dynamics $\mathbf{f}(\mathbf{x}(t), \mathbf{u}(t))$. It should be emphasised that we search for the optimal control policy (an optimal control at each discrete time stage k) producing the minimum cost-to-go over the entire trajectory from the initial to the final stage, and this minimum should be the smallest attainable one (i.e. the global minimum instead of the local minimum).

The main implication from the principle of optimality is that the controls we find optimal from stage k to N are also optimal at a later stage, e.g. from stage $k+1$ to N. It also indicates that the problem is naturally approached in a backwards fashion from the last to the initial stage. For demonstration purposes, consider a simple problem whose solution map is shown in Fig. 4. The system has a single state x, which can take on seven values; the control u can take on three values denoted with the marker coding (circle, star and triangle); and we deal with an N-stage problem of which we show the last three stages and the computed optimal controls corresponding to each state. In the last stage N, we have no optimal controls to compute.

The procedure of computing the optimal controls at stages $N-2$ and $N-1$ of the example map is described in the following. First, let us define the running cost $c_k(x_k, u_k)$ as $L_{p,k}(x_k, u_k) + \ell_k(x_k, u_k)$. Because the final cost is omitted, we begin the procedure at stage $N-1$. In view of Eq. (18), we evaluate the optimal cost-to-go J_{N-1} at a particular state $x_{N-1}^{(1)}$ over the plausible controls $u_{N-1} \in \{u^{(1)}, u^{(2)}, u^{(3)}\}$ as

$$\begin{aligned} J_{N-1}\left(x_{N-1}^{(1)}\right) &= \min\left\{c_{N-1}\left(x_{N-1}^{(1)}, u^{(1)}\right), c_{N-1}\left(x_{N-1}^{(1)}, u^{(2)}\right), \right. \\ &\quad \left. c_{N-1}\left(x_{N-1}^{(1)}, u^{(3)}\right)\right\} \\ &= u^{(2)} \end{aligned} \tag{19}$$

where $u^{(2)}$ is the example optimal control, which minimises the optimal cost-to-go J_{N-1} at the particular state. We store this control and

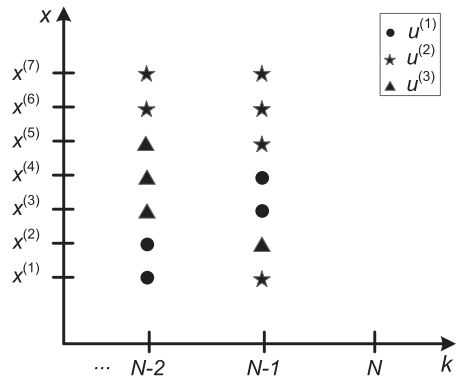


Fig. 4. Example map of marker-coded optimal controls to illustrate the dynamic programming method.

the optimal cost-to-go in the map and repeat this procedure for all of the states.

As we step backwards into stage $N-2$, we utilise the optimal controls and optimal cost-to-go obtained at stage $N-1$ to evaluate the optimal cost-to-go J_{N-2} at each state; for example, for the particular state $x_{N-2}^{(3)}$ we may obtain

$$\begin{aligned} J_{N-2}\left(x_{N-2}^{(3)}\right) &= \min\left\{c_{N-1}\left(x_{N-2}^{(3)},u^{(1)}\right)+J_{N-1}\left(F_k(x_{N-2}^{(3)},u^{(1)})\right),\right. \\ &\quad \left. c_{N-1}\left(x_{N-2}^{(3)},u^{(2)}\right)+J_{N-1}\left(F_k(x_{N-2}^{(3)},u^{(2)})\right),\dots\right\} \\ &= u^{(3)} \end{aligned} \tag{20}$$

where $u^{(3)}$ is the optimal control, which minimises the optimal cost-to-go J_{N-2} at the particular state. In the computation, for example, the solution to J_{N-1} at state $F_k(x_{N-2}^{(3)},u^{(1)})$ is known from computations performed at the previous time stage and we utilise this to evaluate the cost J_{N-2} at state $x_{N-2}^{(3)}$. As we continue with the procedure in this recursive manner to the initial stage, we may resolve the optimal policy of any length between 1 and N stages. The frequently occurring problem is that the next state derived from the state dynamics $F_k(x_k,u_k)$ with some particular control leads to a state unspecified in the state grid and hence the cost-to-go is not evaluated at this state; this problem is solved through the linear interpolation of the cost-to-go.

The DP method can easily deal with those discontinuities arising in hydraulic circuits and help solve optimisation problems arising in complicated hydraulic systems. DP was, for example, used to parametrise hydraulic excavator hybrids in [18]. For more details on the DP method, see [16,17] and [19].

4.1. Proposed modelling approach

DP is a particularly complex approach when dealing with high-dimensional problems such as ours. Indeed, the discrete optimal control problem formulated in Eqs. (16) and (17) is too high-dimensional to solve using the DP algorithm discussed in the previous section. Here we present an improved modelling approach based upon the fact that we should only optimise the movement of redundant DOFs. This reduces our problem dimension from six to a maximum of two states and from three to one control, which enables effective use of the DP algorithm. Considering our emphasis on a typical hydraulic manipulator with three joints, from which one is redundant (Fig. 1), we optimise the movement of this redundant joint. The motion of the other joints can be effectively solved using inverse matrix computations since the remaining system is non-singular.

Let us optimise the motion of the extension joint q_3 of the manipulator. At the velocity level, the simplified system dynamics are hence

$$q_{3,k+1} = q_{3,k} + T_s u_{3,k} \tag{21}$$

where $q_{3,k}$ is the extension joint position at stage k, the control input $u_{3,k}$ at stage k is $\dot{q}_{3,k}$ and T_s is the integration time step. At the acceleration level, a two-dimensional, double-integrator system could be written similarly with the acceleration of the redundant joint $\ddot{q}_{3,k}$ as input. By substituting the optimised position of the redundant joint $q_{3,k}$ into the algebraic solution of the position of the other joints $q_{2,k}$ and $q_{3,k}$, we obtain [7] (see the atan2 version also therein)

$$\begin{aligned} q_{2,k} &= -\operatorname{acos}\left(\frac{n_1}{2L_1\,(L_2+q_{3,k})}\right) \\ q_{1,k} &= \operatorname{asin}\left(\frac{n_2}{n_3}\right) \end{aligned}$$

where

$$\begin{aligned} n_1 &= (r_{x,k}-o_x)^2+(r_{y,k}-o_y)^2-L_1^2-(L_2+q_{3,k})^2 \\ n_2 &= (L_1+(L_2+q_{3,k})\cos(q_{2,k}))(r_{y,k}-o_y) \\ &\quad -(L_2+q_{3,k})\sin(q_{2,k})(r_{x,k}-o_x) \\ n_3 &= L_1^2+(L_2+q_{3,k})^2+2L_1\,(L_2+q_{3,k})\cos\,(q_{2,k}) \end{aligned} \tag{22}$$

and $r_{x,k}$ and $r_{y,k}$, respectively, are the desired Cartesian task-space position samples in the x and y coordinates at time stage k. That is, we also have available discrete samples from the desired Cartesian velocity $\dot{\mathbf{x}}_t$ defined in Eq. (8) at each time stage k. Then we may use Jacobian column vectors $\mathbf{J}_i \in \mathbb{R}^2$ to rewrite Eq. (8) in discretised form as

$$\begin{aligned} \dot{\mathbf{x}}_{t,k} &= \mathbf{J}(\mathbf{q}_k)\dot{\mathbf{q}}_k \\ &= \mathbf{J}_1(\mathbf{q}_k)\dot{q}_{1,k}+\mathbf{J}_2(\mathbf{q}_k)\dot{q}_{2,k}+\mathbf{J}_3(\mathbf{q}_k)u_{3,k} \end{aligned} \tag{23}$$

where the control $u_{3,k}$ is the joint velocity of the extension joint coordinate $\dot{q}_{3,k}$. From Eq. (23), we solve the desired Cartesian velocity with the contribution of the optimised extension joint included so

$$\begin{aligned} \dot{\mathbf{x}}_{t_r,k} &= \dot{\mathbf{x}}_{t,k}-\mathbf{J}_3(\mathbf{q}_k)u_{3,k} \\ &= \mathbf{J}_1(\mathbf{q}_k)\dot{q}_{1,k}+\mathbf{J}_2(\mathbf{q}_k)\dot{q}_{2,k} \end{aligned} \tag{24}$$

where $\dot{\mathbf{x}}_{t_r,k}$ contains the Cartesian velocity required from joints q_1 and q_2 to maintain the desired Cartesian trajectory $\dot{\mathbf{x}}_{t,k}$. Because the remaining system is non-singular, we solve the angular joint velocities $\dot{q}_{1,k}$ and $\dot{q}_{2,k}$ in a straight forward manner using matrix inversion:

$$\begin{bmatrix}\dot{q}_{1,k}\\ \dot{q}_{2,k}\end{bmatrix} = \begin{bmatrix}\mathbf{J}_1(\mathbf{q}_k) & \mathbf{J}_2(\mathbf{q}_k)\end{bmatrix}^{-1}\dot{\mathbf{x}}_{t_r,k} \tag{25}$$

where the Jacobian inverse matrix $[\mathbf{J}_1(\mathbf{q}_k)\,\mathbf{J}_2(\mathbf{q}_k)]^{-1}$ is symbolically computed in advance.

Extending the formulation to the acceleration level, we replace Eq. (21) with a standard two-dimensional double-integrator system with $\ddot{q}_{3,k}$ as the control $u_{3,k}$. Then, as before, we solve the unknown joint positions and velocities with Eqs. (22)–(25) and solve the non-singular system to obtain the joint accelerations

$$\begin{bmatrix}\ddot{q}_{1,k}\\ \ddot{q}_{2,k}\end{bmatrix} = \begin{bmatrix}\mathbf{J}_1(\mathbf{q}_k) & \mathbf{J}_2(\mathbf{q}_k)\end{bmatrix}^{-1}\times\left(\ddot{\mathbf{x}}_{t,k}-\dot{\mathbf{J}}(\mathbf{q}_k,\dot{\mathbf{q}}_k)\dot{\mathbf{q}}_k-\mathbf{J}_3(\mathbf{q}_k)u_{3,k}\right) \tag{26}$$

where the subtraction on the right-hand side in parenthesis contains the Cartesian acceleration required from joints q_1 and q_2 to maintain the desired Cartesian trajectory $\ddot{\mathbf{x}}_{t,k}$ and the Jacobian time derivative $\dot{\mathbf{J}}(\mathbf{q},\dot{\mathbf{q}})$ is symbolically computed in advance using the well-known chain rule of differentiation. In the above, we assumed that the desired Cartesian trajectory is twice-differentiable and sampled.

A similar approach formulated with torque input can be found in [20], but the approach here presents a simpler and more general system. In [21], the solution obtained in [20] was criticised for being a complex formulation of the acceleration, whereas our velocity level one-dimensional approach is merely a function of the joint positions and velocities. This considerably simplifies the redundancy resolution and yields the global solution for the CP system with relative ease. Furthermore, our two-dimensional acceleration level approach, which is not formulated from the perspective of manipulator dynamics, preserves simplicity over [20] and facilitates the introduction of general cost functions dependent on the manipulator motion state. The major benefit of the acceleration level solution over the velocity level solution is that the acceleration level solution satisfies physical joint acceleration limits. For simplicity, the highly complex pump

equations and actuator pressure dynamics are omitted from the system dynamics, which means that we assume the pump can respond to the flow required by the actuators and the hydraulic fluid is incompressible. The problem formulation presented here means that a standard DP solution is feasible with the proposed approach.

5. Proposed cost functions

In hydraulic manipulators, the energy consumption of the actuators does not equal the energy consumed by the hydraulic components because of the pressure losses over the control valves when the actuators are subject to unequal loads. Thus, to effectively solve the redundancy resolution problem, we propose cost functions formulated from the standpoint of the hydraulic system instead of the manipulator dynamics or actuators.

5.1. Pump flow rate

Minimising the pump flow rate over the Cartesian trajectory decreases pumping effort and minimises the hydraulic energy consumption of the CP system. The minimisation is plausible because of the controllability of the variable displacement pump. This minimisation has particular relevance when the hydraulic cylinders have different sizes, thus leading to varying cylinder flow requirements. Due to the high variability of cylinder sizes in practice, the problem of finding joint trajectories of the least pump flow over the Cartesian path is tractable. The different piston and piston rod areas in single-rod cylinders extend the optimisation potential.

The discontinuity and nonlinearity of the pump flow cost function, which occurs because of the variation of displaced area as a function of the direction of motion, is not a problem in standard DP. In view of Eq. (16), the pump flow rate cost function such that $Q_{\mathrm{p},k} \geq 0$ may be written at time stage k by

$$\begin{aligned} L_{\mathrm{p},k} &= Q_{\mathrm{p},k} \\ &= \sum_{i=1}^{3} \{Q_{\mathrm{A}i,k} - Q_{\mathrm{B}i,k}\} \\ &= \sum_{i=1}^{3} \{v_{i,k}\,[A_{\mathrm{A}i} H(v_{i,k}) - A_{\mathrm{B}i} H(-v_{i,k})]\} \\ &= \sum_{i=1}^{3} \{\dot{q}_{i,k} r_{\mathrm{n}_i,k}\,[A_{\mathrm{A}i} H(\dot{q}_{i,k}) - A_{\mathrm{B}i} H(-\dot{q}_{i,k})]\} \end{aligned} \tag{27}$$

where $Q_{\mathrm{A}i,k}$ is the flow rate to the piston side of the cylinder i, $Q_{\mathrm{B}i,k}$ is the flow rate to the piston rod-side of the cylinder i, $H(\dot{q}_{i,k})$ is the piecewise Heaviside step function, $v_{i,k}$ denotes the cylinder velocity of actuator i, the discontinuous cylinder area is denoted with $A_{\mathrm{A}i}H(\dot{q}_{i,k}) - A_{\mathrm{B}i}H(-\dot{q}_{i,k})$, in which $A_{\mathrm{A}i}$ is the piston area and $A_{\mathrm{B}i}$ is the piston rod-side area of cylinder i, and $r_{\mathrm{n}_i,k}$ is the torque arm of cylinder i. For simplicity, the hydraulic fluid is assumed to be incompressible. Using a cylinder differential connection would change the piston side area $A_{\mathrm{A}i}$ to $A_{\mathrm{A}i} - A_{\mathrm{B}i}$ in the case of cylinder extension. This change would mean that the pump flow rate requirement for this particular movement reduces to $v_{i,k}\,(A_{\mathrm{A}i} - A_{\mathrm{B}i})$ or $v_{i,k}\,A_{\mathrm{r}i}$, where $A_{\mathrm{r}i}$ is the circular area of the piston rod of cylinder i.

The Heaviside step function is defined by

$$H(\dot{q}_{i,k}) = \begin{cases} 0 & \text{if} \quad \dot{q}_{i,k} < 0 \\ \frac{1}{2} & \text{if} \quad \dot{q}_{i,k} = 0 \\ 1 & \text{if} \quad \dot{q}_{i,k} > 0 \end{cases} \tag{28}$$

where the intermediate value at zero joint velocity is defined as half for convenience. This discontinuous Heaviside step function may be rewritten in differentiable form after some mathematical manipulations:

$$H(\dot{q}_{i,k}) \approx \hat{H}(\dot{q}_{i,k}) = \frac{1}{1 + e^{-2s\dot{q}_{i,k}}} \tag{29}$$

where s is a real positive number. By using this approximation, the flow rate objective is differentiable for all real $\dot{q}_{i,k}$ and may be convenient in practice.

In Eq. (27), the $A_{\mathrm{B}i}$ area is negatively signed to enforce positive $Q_{\mathrm{B},k}$ when the joint velocities are negative. We used the properties of tangential velocity, which states that $v_{i,k}$ can be written as $\dot{q}_{i,k} r_{\mathrm{n}_i,k}$, and also the property, which states that $H(v_{i,k})$ can be rewritten as $H(\dot{q}_{i,k})$ since $r_{\mathrm{n}_i,k}$ is positive. The torque arm $r_{\mathrm{n}_i,k}$ at time stage k is defined by

$$r_{\mathrm{n}_i,k} = \frac{L_{i_1} L_{i_2} \sin\left(q_{i,k} + q_i^{(0)}\right)}{\sqrt{L_{i_1} + L_{i_2} - 2L_{i_1} L_{i_2} \cos\left(q_{i,k} + q_i^{(0)}\right)}} \tag{30}$$

if $q_{i,k}$ is rotational, otherwise $r_{\mathrm{n}_i,k}$ is one. The distances L_{i_1} and L_{i_2} are the constant distances between the centre of rotation and lower and upper cylinder joints, respectively. The initial value on the torque arm corresponds to a fully retracted piston when the joint coordinate $q_{i,k}$ is at its minimum value $q_{i_{\min}}$. Hence, $q_i^{(0)}$ is denoted by $q_{i_{\mathrm{c}}} - q_{i_{\min}}$. The quantity $q_{i_{\mathrm{c}}}$ is the angle of the triangle opposite a fully retracted cylinder. The triangle is formed by the upper cylinder joint, lower cylinder joint and rotational joint $q_{i,k}$.

5.2. Constant supply pressure system: energy consumption

By optimising the movement of the redundant joint to minimise the flow delivered to the actuators, the energy consumption of the CP system is minimised. We see this from the hydraulically produced power of the constant pressure system, written in view of Eq. (16)

$$L_{\mathrm{p},k} = \frac{p_{\mathrm{s}} Q_{\mathrm{p},k}}{\eta_{\mathrm{t},k}} \tag{31}$$

where p_{s} is the constant supply pressure, $Q_{\mathrm{p},k}$ is the pump generated flow rate defined in Eq. (27) at time stage k and $\eta_{\mathrm{t},k}$ is the total energy efficiency of the hydraulic pump and driving motor at time stage k. Because weighting $L_{\mathrm{p},k}$ with $1/p_{\mathrm{s}}$ yields the pump flow rate cost, the objectives are ideally equal. This holds the assumption that the constant pressure has been fixed pre-optimisation. Moreover, based on our assumption that the hydraulic energy cannot be reused, the energy consumed maintains positivity. The simplified one-dimensional system model previously presented can be used with this cost function, but the joint acceleration limits may not be obeyed. The total energy efficiency depends on the prevalent operating point, i.e. the supply pressure, displacement and rotation speed of the pump, but to showcase the basic capability of the redundancy resolution and due to a lack of realistic efficiency data, we assume a constant energy efficiency in our investigations.

5.3. Load-sensing system: energy consumption

Systems based on LS architecture generate pressure losses over the control valves whenever the loading between actuators is unequal because of the system's nature, in which the highest actuator pressure is demanded at the pump level. Because the supply pressure varies depending on the load, instead of minimising flow,

we must specifically minimise the hydraulic power produced by the LS system

$$L_{p,k} = \frac{p_{s,k} Q_{p,k}}{\eta_{t,k}} \tag{32}$$

where $p_{s,k}$ is the supply pressure at time stage k, $Q_{p,k}$ is the pump flow rate at time stage k defined by Eq. (27) and $\eta_{t,k}$ is the total efficiency of the hydraulic pump and driving motor at time stage k.

Omitting the dynamics of the variable displacement pump, the supply pressure $p_{s,k}$ varies in unison with the highest actuator pressure

$$p_{s,k} = \max\{p_{1,k}, p_{2,k}, p_{3,k}\} + \Delta p_{LS} \tag{33}$$

where $p_{1,k}$ denotes the chamber pressure of the lift cylinder, $p_{2,k}$ denotes the chamber pressure of the tilt cylinder, $p_{3,k}$ denotes the chamber pressure of the extension cylinder and Δp_{LS} is the LS pressure margin, which is conventionally set to approximately 2 MPa. The cylinder chamber pressure of actuator i is solved from

$$p_{i,k} = \left(\frac{|F_{i,k}|}{A_{i,k}} + p_{BP,i,k}\right) H(F_{i,k} v_{i,k}) \tag{34}$$

where the actuator force $F_{i,k}$ at time stage k can be computed using $T_{i,k}/r_{n_i,k}$, with $r_{n_i,k}$ defined by Eq. (30), $p_{BP,i,k}$ is the positive cylinder back-pressure at time stage k and piston area $A_{i,k}$ is denoted with $A_{Ai}H(\dot{q}_{i,k}) + A_{Bi}H(-\dot{q}_{i,k})$, which varies as a function of the direction of motion. We see that $H(F_{i,k}v_{i,k})$ sets the required actuator pressure to zero when dealing with negative energy, and the absolute value on $F_{i,k}$ satisfies the requirement for non-negative actuator pressure.

The actuator forces $F_{i,k}$ and torques $T_{i,k}$ are solved from manipulator dynamics, which can be formulated based on well-known Lagrangian or robotic conventions. Centres of the mass positions of the links r_i and the link masses m_i in Fig. 5 are provided in Appendix A. Using these parameters, we obtain the inertia matrix and gravitational component using the procedure described in [7].

The back-pressure of cylinder i can be written as

$$p_{BP,i} = \frac{A_{B,i}}{A_{A,i}} p_{BP,B,i} H(v_i) + \frac{A_{A,i}}{A_{B,i}} p_{BP,A,i} H(-v_i) \tag{35}$$

where time indices are omitted for clarity. Back-pressure is taken from the rod-side when the cylinder extends by using the Heaviside function. Similarly, back-pressure is taken from the piston-side when the cylinder retracts. The cylinder area ratios scale the back-pressure from rod to piston-side and vice versa. In the case of a pressure-compensated hydraulic valve, the back-pressures of cylinder chambers A and B, i.e. $p_{BP,A,i}$ and $p_{BP,B,i}$, respectively, can be estimated in the steady-states by the multiplication of the valve's constant pressure difference with a certain coefficient [22]. When the back-pressures are insignificant, like in our closed-loop controlled separate meter-in and separate meter-out orifice configuration, these steady-state back-pressure functions can be omitted.

The maximum function in Eq. (33) is discontinuous but could be approximated with a differentiable expression for the sake of practical implementation after some mathematical manipulations:

$$\max\{p_{1,k}, p_{2,k}, p_{3,k}\} = n_p/4 \tag{36}$$

where n_p is given by $p_{1,k}+p_{2,k}+|p_{1,k}-p_{2,k}|+2p_{3,k}+|p_{1,k}+p_{2,k}+|p_{1,k}-p_{2,k}|-2p_{3,k}|$ and the absolute values $|p_{i,k}|$ should be approximated with the square root of ${p_{i,k}}^2 + \epsilon$. This approximation originates from the well-known definition of the maximum function of two variables via absolute values. The accuracy of this approximation improves when we decrease the value of the real positive ϵ (e.g. to class 10^{-6}).

6. Numerical examples

The purpose of these numerical examples is to compare the global and local solutions in relation to typical hydraulic manipulator applications and to showcase the superior performance of the global approach, which is due to the properties of the optimal control problem. We are particularly interested in showing how much energy can be saved in relation to the cost functions. We perform a comparison and discuss some parameters' effect on the solutions.

6.1. Setting up the numerical examples

Let us define our manipulator using the parameters supplied in Appendix A. By applying these, we obtain the workspace shown in Fig. 6. The reachable workspace without the extension joint ($q_3 = 0$) is shown with a circular marker. The load mass was fixed at 475 kg, which is a reasonable choice considering the heavy-duty lifting carried out at construction sites. This load mass was also close to the load capacity of the system. The same weight was employed to closed-loop control tests in [1]. We setup our comparison so a variety of optimisation methods complete a fictitious task cycle in which the end-effector is driven through triangular paths comprising diagonal, vertical and horizontal path, completed in this order. Throughout the

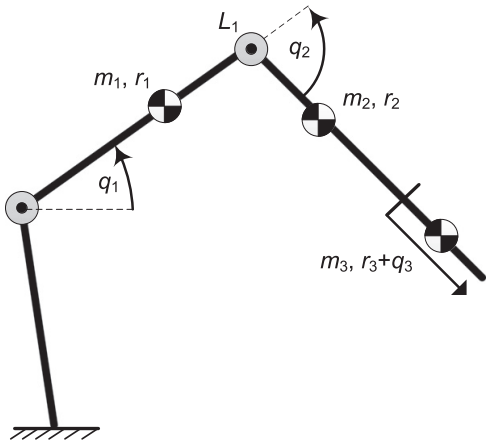


Fig. 5. Positions of the centre of masses of the links.

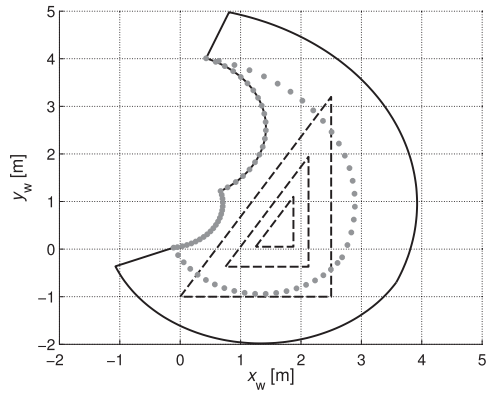


Fig. 6. Test cycles for analysing attainable energy-saving in the manipulator workspace.

experiments, the Cartesian paths between two points are generated with a quintic rest-to-rest polynomial trajectory, which provides smooth positional references for deriving the desired velocity and acceleration trajectories [23].

On the hydraulic side, the joint-actuating cylinders were sized as follows: ⌀80/45 − 0.545, ⌀80/45 − 0.545 and ⌀50/30 − 1.04, respectively, for the first (lift function), second (tilt) and third (extension) joints. These parameters were derived from a commercial construction crane. The cylinder velocities and accelerations were conservatively limited to the values shown in Appendix A. In addition, we briefly investigate the parameter sensitivity's effect on the redundancy resolution.

The DP approaches with various cost functions are compared with non-conventional pseudo-inverse approaches formulated in actuator coordinates, and the fmincon function from the Optimization toolbox in Matlab and the DIDO application package [24,25] are compared with the application to the described manipulator. The pseudo-inverse applied was the joint-limited (joint ranges and joint velocity limits satisfied) null-space projection method introduced by Flacco et al. [26], but we extended it to the actuator space so norm $\mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v}$ as the weighted or norm $\mathbf{v}^{\mathrm{T}}\mathbf{v}$ as the unweighted version are minimised instead of the standard norm $\dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{q}}$. The joint ranges (cylinder strokes) and velocity limits as well as the reference trajectory were satisfied. The weighting matrix is a diagonal matrix $\mathbf{A} \in \mathbb{R}^{3\times3}$ in which the diagonal terms A_{ii} are given by $A_{\mathrm{A}i}H(v_i) + A_{\mathrm{B}i}H(-v_i)$. The Matlab function fmincon (ver. 2013b) was set up with the default interior-point method; the system used was Eqs. (11)–(15) integrated with a fixed-step Runge-Kutta solver. For the fmincon algorithm, we used the high-dimensional system because the solution resulting from the proposed low-dimensional system more frequently failed to converge while satisfying the constraints; whereas for the DIDO software, we successfully used the low-dimensional system. It is a well-known fact that pseudospectral methods (that is implemented for example in DIDO software) can yield satisfactory trajectories with relatively few discretisation nodes, albeit the value of the cost may then be imprecise. To improve the cost of the DIDO solution and smoothness of DIDO trajectories, the DIDO solver was run with higher nodes via a bootstrapping approach. In this approach, regarding the CP case, the solver was first run with a 20 discretisation node solution that was inputted as a guess for a 30 node solution that, in turn, was inputted as a guess for a 60 node solution that was finally inputted as a guess for a 90 node solution. The LS case was treated similarly, but using a higher number of nodes. We noticed that a DIDO solution is obtained significantly faster when using a continuous form of the flow rate cost function. Thus, we employed the continuous Heaviside approximation. The position of the extension joint was initialised with the smallest feasible value satisfying the other joint ranges, and the remaining joint positions were solved using Eq. (22) in all of the methods. The initial and final joint velocities were set to zero in the optimal control methods.

Energy savings were firstly computed using a simplified model which did not possess actuator pressure dynamics, friction effects or LS pump pressure dynamics. Energy savings were secondly computed using a full-scale, closed-loop simulation of the hydraulic system to demonstrate that our simplified problem formulation is pragmatic and the neglected hydraulic aspects of the problem could be omitted. Actuator pressure dynamics, friction effects and pump pressure dynamics were included in the full simulation model. Energy saving results are presented for both the simplified and full simulation case in the following sections. In the closed-loop simulation case, hydraulic cylinder chamber pressures were controlled independently using pressure-compensated valves, which were set-up in a separate meter-in and separate meter-out orifice configuration. The separate meter-in and meter-out valves were controlled so that valve notch connections P-A & B-T and P-B & A-T were simultaneously opened (see Figs. 2 and 3). The notches' magnitude of opening was of course otherwise controlled independently. Hydraulic system parameters were set to the same values in each closed-loop simulation to yield comparable results. The adaptive robust control approach [27,28] was used as the motion controller; however, the cross-port valve was excluded. The adaptive robust controller was implemented for each manipulator function, allowing the controller's robustness to modelling uncertainties to dominate neglected joint coupling terms. To avoid cavitation, the so-called offside cylinder chamber pressures during the motion were regulated to a constant 1 MPa.

6.2. Constant-pressure system

The cost function used for the comparison in Table 2 is Eq. (16), into which the pump flow cost from Eq. (27) is substituted. Therefore, the cost function is the sum of the pump flow rate over the Cartesian trajectory in which the largest triangular path is driven with 10 seconds spent on each edge. The cost function values are scaled with the minimum, i.e. best, result. We searched for the global solution from the infinite space of solutions without any restriction apart from the constraints on joint ranges, joint velocity limits and joint acceleration limits set in Appendix A. The maximum supply pressure level was not restricted in any way. We obtained closed-loop tracking performances comparable to [27,28] when using the optimised joint motion trajectories as the motion controller's references, i.e. the position tracking errors were generally less than 0.01 rad or 0.01 m.

The energy savings predicted by the simplified model, which neglected conventional hydraulic system dynamics, and the more accurate closed-loop simulation prediction, which included these effects, are mostly comparable, showing that the degree of model simplification in the problem formulation is justified. The DP approach with the CP cost function yields the lowest cost in both cases, although the DIDO approach set to minimise Eq. (27) has an almost identical cost to the DP approach. Still, keeping that DP approach as our baseline, a comparison shows that the DP methods written to minimise the LS cost in view of Eq. (32), the positive actuator energy cost or the actuator velocity cost $\mathbf{v}^{\mathrm{T}}\mathbf{v}$ yield sub-optimal joint trajectories that require a somewhat higher flow to perform the test cycle. The fmincon function yields a very poor local optimum that is significantly inferior to the global optimum. The weighted pseudo-inverse yields the worst solution, in which the pump flow over the test cycle was over 30% greater. Surprisingly, the unweighted pseudo-inverse written in the cylinder coordinates yields a decent result. All of the results concern the largest triangle path in Fig. 6 completed in 30 s, or 10 s per edge, which amounts to a Cartesian velocity of the end-effector of roughly 0.5 m/s on the diagonal path.

In reference to the smaller triangular paths, we saw the energy-saving potential clearly decreasing because of the smaller area covered, i.e. optimisation potential is lost with a decreasing range of motion. The DP solutions owing to the discretisation are prone to having slightly jagged edges, but proper filtering smooths these out.

Table 2
Constant pressure system: comparison of optimal and sub-optimal approaches (with joint limits) based on energy savings predicted by the simplified model and closed-loop simulated full-scale model.

Method (minimised cost)	Simplified rel. cost function	Full rel. cost function
DP (CP)	1.000	1.000
DIDO (CP)	1.003	1.002
DP (LS)	1.062	1.060
DP (Act. velocity)	1.087	1.086
DP (Pos. actuator)	1.098	1.098
Act. p-inv. (Unweighted)	1.127	1.124
fmincon (CP)	1.166	1.163
Act. p-inv. (Weighted)	1.308	1.308

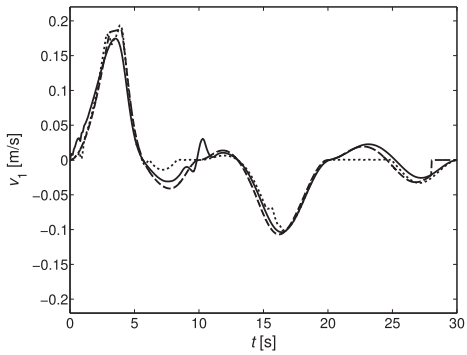
The cost function values in the tables throughout this paper were computed after this low-pass filtering. The low-pass filtered (cutoff frequency 5 Hz) cylinder velocity trajectories are in Fig. 7a–c, which show the largest test cycle.

The proposed simplified first-order DP approach was parametrised by state and control grid size, with the state being divided into 200 discrete values and the control being divided into 101 discrete values. In addition, the algorithm was highly computationally efficient and the modelling accuracy (with its discretisation) was more than adequate. The time was discretised into 0.05 second steps. Surprisingly, this enabled a significantly faster solution and lower system memory consumption than the fmincon method, which uses the interior-point algorithm. As seen, the fmincon can very easily yield a local minimum to our problem. Therefore, global optimal control is seen as a highly attractive solution.

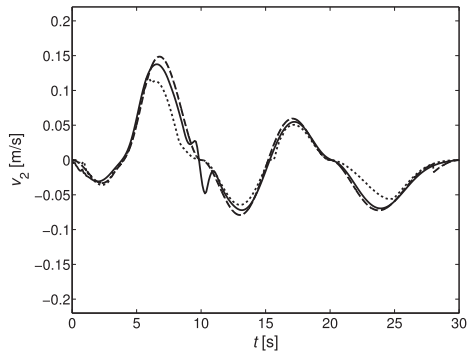
The second-order model, however, allows the inclusion of joint acceleration limits we applied to our comparison. DP with the joint acceleration constraints intuitively yields a cost that is always higher or the same as the cost obtained from DP without joint acceleration constraints. In terms of discretisation, the state space of the extension cylinder position was divided into 125 discrete values, the extension cylinder velocity into 101 values and the extension cylinder acceleration into 201 values. DP with the second-order model is still faster than the fmincon algorithm even though the increase in the number of states leads to an exponential increase in the computational burden. Decreasing the grid sizes in this case reduces the computational effort without any significant sacrifice in performance. Halving the state grid size halves the requirement on system memory and computation time, and the same applies for the control grid. The second-order approach was clearly not as computationally efficient as the first-order approach, but decreasing the grid sizes even more is plausible if faster execution is desired, albeit at the expense of inferior trajectories. The DIDO approach possessed a lower computational complexity than the DP approach and a somewhat lower memory requirement.

Fig. 8 demonstrates the convergence properties of the dynamic programming algorithm in the CP case. It shows that even the coarser state and control grids are capable of producing solutions that are in reasonably close agreement with the global optimum. The anticipated difference between the coarser and denser grid solutions is the smoother motion profiles of the denser solutions. The discretisation used corresponds to the second to last marker in Fig. 8. Finally, as a major convenience, the problem concerning the CP system may be solved without any knowledge about the manipulator dynamics.

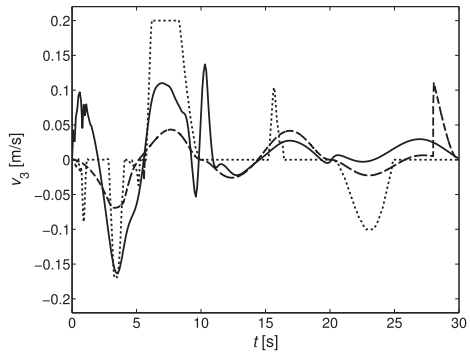
The energy saved can be increased by decreasing the cylinder areas since the actuator flows are reduced. In some cases, however, the decrease in cylinder area leads to an intolerable increase in the supply pressure level. Even minor realistic cylinder area changes can increase the energy savings and therefore extend the optimisation potential. Consider our numerical example in which the sizes of the tilt and lift actuators are the same. If we changed the tilt cylinder size (joint 2) from ⌀80/45 − 0.545 to ⌀80/56 − 0.545, the flow requirement on the pump decreased by 5% over the largest triangular path. The reduction was 4% in the medium and small triangular paths with



(a) Lift cylinder velocity v_1 (m/s) over time (s)



(b) Tilt cylinder velocity v_2 (m/s) over time (s)



(c) Ext. cylinder velocity v_3 (m/s) over time (s)

Fig. 7. CP optimal dynamic programming trajectory (- - - -), suboptimal fmincon trajectory (—) and pseudo-inverse trajectory (– –).

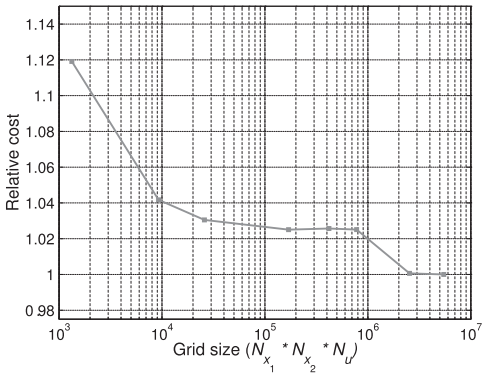


Fig. 8. Convergence of the dynamic programming algorithm in the constant-pressure case; the scale for the x-axis is logarithmic to improve readability, and the y-axis cost function value is scaled with the cost obtained using the densest grid size.

this moderate change. But since the maximum supply pressure in the cycle increased by well over 5%, the change was clearly not advisable. If we changed the extension cylinder size from ⌀50/30 − 1.04 to ⌀45/30−1.04, the energy saved on the largest triangular path was 4%, with negligible effect on the supply pressure level. It may be that the extension cylinder of the typical hydraulic manipulator has the most potential for size optimisation. Component size optimisation is, however, not the main focus of this paper, but even this example demonstrates the energy savings obtained by sizing the hydraulic cylinders used in a work cycle.

6.3. Load-sensing system

As before, the cost function used for the comparison in Table 3 is Eq. (16), into which the LS cost from Eq. (32) is substituted. Therefore, the cost function is the sum of consumed LS energy over the Cartesian trajectory in which the largest triangular path is driven, with 10 s spent on each edge. The cost function values are scaled with the minimum result. The same state and control grids were applied as in the CP case, and the second-order model with joint ranges, cylinder velocity and acceleration limits were used. The closed-loop control performances in the simulations, in which the LS pump's time constant was set to 0.20 s [29], were similar to the CP case.

The energy savings predicted by the simplified model and closed-loop simulation are somewhat different compared to the CP case. The results imply that the energy savings obtainable in an actual LS system are lower than predicted. Still, the DP approach minimising the LS energy consumed yielded the best solution. We kept this as our baseline for comparison. The weighted pseudo-inverse formulated in the actuator coordinates produces a cost function value, which was almost 30% higher than that of the best DP solution in the simulations. The fmincon also produces a decent result, particularly in the simulated case. The computational demand of the fmincon algorithm is notably high. The pseudo-inverses produce inferior results, because they are only point-wise optimal and do not satisfy joint acceleration limits. The DP methods produce significantly better trajectories than the point-wise methods considering the moderate optimisation potential for the manipulator with one redundant joint. The DP approach minimising the CP energy consumed yields a cost which is reasonably close to the global optimum in the closed-loop simulation. As before, the DIDO solver produces a satisfactory trajectory, particularly for closed-loop simulation.

Table 3
Load-sensing system: comparison of optimal and sub-optimal approaches (With inertial and gravitational effects and joint limits) based on energy savings predicted by the simplified model and closed-loop simulated full-scale model.

Method (minimised cost)	Simplified rel. cost function	Full rel. cost function
DP (LS)	1.000	1.000
DIDO (LS)	1.087	1.012
DP (Pos. actuator)	1.109	1.027
fmincon (LS)	1.154	1.050
DP (CP)	1.201	1.083
DP (Act. velocity)	1.247	1.135
Act. p-inv. (Unweighted)	1.259	1.154
Act. p-inv. (Weighted)	1.413	1.293

The DP method which penalises positive actuator work is nearly optimal even from the LS perspective. The numerical discrepancy in closed-loop simulation is minor, but the desired trajectories differ, as shown in the cylinder velocity trajectories below in Fig. 9a–c. This discrepancy originates from the LS pressure losses which arise from the requirement of highest actuator pressure at the pump level. When these optimised joint trajectories are driven in closed-loop and compared, some of the optimisation potential of the LS system is lost because of the pump's non-ideal pressure dynamics. As before, the energy-saving potential decreased when the size of triangular path was reduced because of the smaller area covered.

The desired pump supply pressure (LS pressure) based on the simplified model is shown in Fig. 9d as the joints were driven through the optimal joint trajectories. This figure illustrates what the LS pressure would be with the other DP solutions for a comparison. The LS optimal solution clearly searches for a lower supply pressure than does the CP optimal solution. The actuator energy optimal solution also has a higher LS pressure demand in general than did the LS solution. The maximum actuator pressure at the pump in LS systems affects the consumption of all the actuators, which is why the LS solution must search for a path of lower pressure.

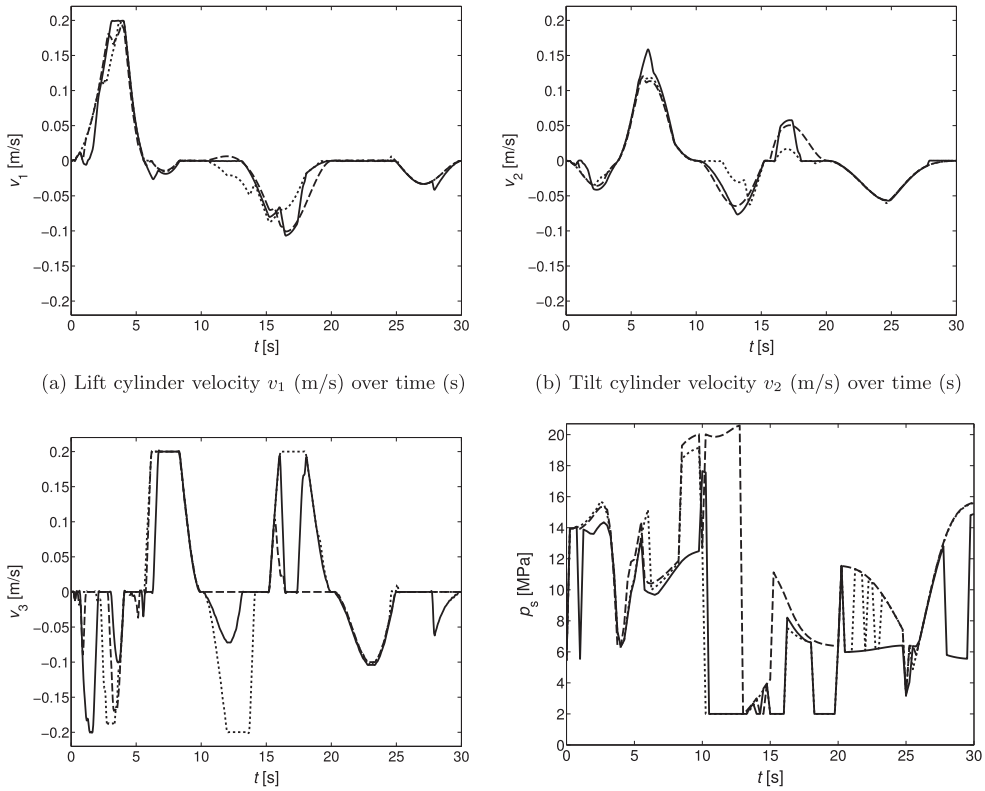
Fig. 10 demonstrates the convergence properties of the dynamic programming algorithm in the LS case. Compared to the CP case, the coarser state and control grids produce solutions that are further from the global optimum. True global optimality requires a reasonably dense grid size. The discretisation we used corresponds to the second to last marker in Fig. 10.

Changing the tilt cylinder size (joint 2) from ⌀80/45 − 0.545 to ⌀80/56 − 0.545 increased the LS energy required over the largest triangle trajectory by almost 6%, even though the flow required by the actuators was reduced. A significant increase in LS pressure explains this. Changing the extension cylinder size from ⌀50/30 − 1.04 to ⌀45/30 − 1.04 was more favourable because the pump flow was reduced while the highest actuator pressure was mostly unaffected. The modification in area decreased the LS energy required by 4%. It is essential to note that while the extension cylinder pressure increased, the LS pressure for the most part did not.

Finally, the simplified and the more complete closed-loop simulation model are compared in Fig. 11a–b. The pump flow rate predictions of the models are consistent, although fluid compressibility was only included in the closed-loop simulation model. The supply pressure predictions are understandably less consistent because approximately 1 MPa of the supply pressure offset is due to the cylinder back-pressures omitted from the simplified model. Pump pressure dynamics and closed-loop control behaviour account for the remaining offset seen in the supply pressure comparison.

7. Discussion

In the literature and research papers thus far, the redundancy resolution problem has been resolved at the actuator level. Our results have shown this could be a suboptimal approach at the hydraulic system level. Moreover, the hydraulic redundancy resolution problem



(a) Lift cylinder velocity v_1 (m/s) over time (s)

(b) Tilt cylinder velocity v_2 (m/s) over time (s)

(c) Ext. cylinder velocity v_3 (m/s) over time (s)

(d) Estimated LS pressure p_s (MPa) over time (s)

Fig. 9. LS optimal dynamic programming trajectory (—), CP optimal dynamic programming trajectory (– –) and positive actuator energy optimal dynamic programming trajectory (- - - -).

is a complex problem that cannot necessarily be resolved to global optimality using standard direct optimisation tools. Our results show that the local optimums obtained can be very poor and are close to point-wise optimal trajectories at times.

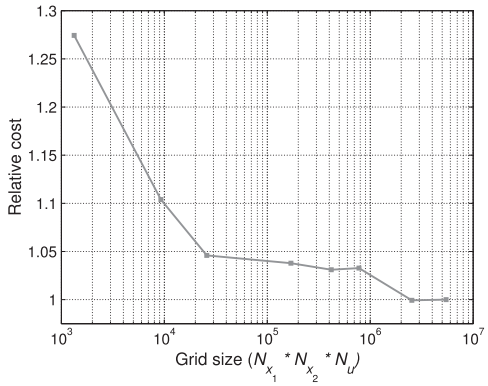
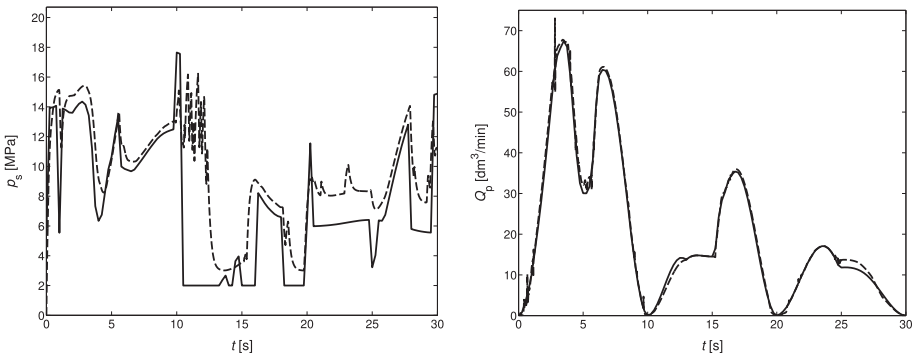


Fig. 10. Convergence of the dynamic programming algorithm in the load-sensing case; the scale for the x-axis is logarithmic to improve readability, and the y-axis cost function value is scaled with the cost obtained using the densest grid size.

Moreover, the energy saved, as demonstrated using the DP and DIDO approaches, can be very significant (e.g. 30% compared to some point-wise optimal methods). These results demonstrate the reduction attainable in the energy consumption of a typical hydraulic manipulator which has one redundant joint. Therefore, the energy savings obtained cannot be exceptionally high. The numerical examples presented here are formulated using realistic parameters, so the results can be considered practically plausible. The extra reach by the extension cylinder in the numerical example was roughly 1 m, where as in some other applications the reach may be over 2 m. This could have some significance when determining the optimisation potential in other similar manipulators.

Some simplifications have been made in the problem formulation to ease the computational burden: (1) the ideal hydraulic fluid was assumed to be incompressible and (2) the pump dynamics were neglected. The first assumption seems reasonable because the amount of compressed fluid delivered by the pump should be relatively low compared to the incompressed fluid leaving the pump. This observation was verified because energy savings predicted by the simplified model had results comparable to the closed-loop simulations in the CP case. The second assumption means that we assumed that the pump responds to the commanded supply pressure within the discretised time interval of 0.05 s. This simplification, which was made for computational convenience, affects the validity of the results from a real manipulator because an actual pump cannot respond this fast. Particularly, we saw that including pump



(a) LS case: Supply pressure p_s (MPa) over time (s) (b) CP case: Pump flow rate Q_p (dm^3/s) over time (s)

Fig. 11. Model comparison: simplified model used in optimisation (—) and full model used in closed-loop simulations (– –).

dynamics in a closed-loop simulation of the optimum joint trajectories, some energy saving potential in the LS case is lost. However, the simplification could still be justified because it decreases computational complexity. Moreover, the DP approach minimising the CP energy consumed, formulated without the knowledge of manipulator dynamics, was reasonably accurate in decreasing the LS case cost.

In this study we assumed a separate meter-in and separate meter-out valve configuration (SMISMO). This SMISMO configuration could allow for additional savings from differential connection utilisation, but we decided to omit valve level optimisation from this paper and focus on the redundancy resolution through cylinder motion optimisation that is the core of the paper. A differential connection to the hydraulic cylinder was thus not utilised in the simulations. At the moment this pump flow rate reducing connection has also not been used in the real hydraulic manipulator, which was the basis for the simulations. Cylinder back-pressures were also omitted from the optimisation because the back-pressures could be regulated to a minor 1 MPa value in the closed-loop simulations using the SMISMO valve configuration. Furthermore, a fixed hydraulic pump efficiency was assumed in the simulations. Considering a variable pump efficiency could have influenced the simulation results and changed the outcome of this study.

Our approach is intended for offline optimisation, and the results are seen to motivate energy optimal redundancy resolution at the hydraulic level in future applications aiming to automate some construction crane work cycles. When the automation comes to fruition, the optimal joint trajectories could computed beforehand and retrieved from a database when needed. Online application of our solution is not possible when the workspace trajectory is unknown. To enable online application, one would need to predict the future trajectory over a short time horizon or design an improved point-wise optimal approach: the pseudo-inverse one used for comparison in this paper is clearly suboptimal. We assume that a point-wise optimal solution with appropriate adaptive weights could yield a satisfactory result, but finding the weights would be difficult.

Cylinder optimisation has some significance because it can further improve the redundancy resolution at the hydraulic level. In CP systems, this cylinder optimisation is as simple as decreasing the cylinder areas, which improves the energy savings if the pump supply pressure does not need to be boosted. In LS systems, we found the extension cylinder area to be most often reducible without increasing the LS pressure level, which leads to decreased LS energy requirement. However, on an LS system, cylinder area optimisation has intricate ties to the load mass and geometry of the manipulator. The effect of varying load mass was not evaluated, but it may be assumed that the load affects the energy savings obtained via redundancy resolution.

In our work, we did not consider energy recuperation because the energy recuperation and hybrid systems linked to these cases are still in their infancy on an application level. In addition, we focused on traditional systems because the energy optimisation at the hydraulic level had not yet been studied. In view of systems with energy recuperation, the pump flow rate cost function should be written, omitting time indices, as

$$Q_p = \dot{q}_1 r_{n_1}(A_{A1} - A_{B1}) + \dot{q}_2 r_{n_2}(A_{A2} - A_{B2}) + \dot{q}_3(A_{A3} - A_{B3}) \tag{37}$$

where r_{n_i} denotes the torque arms, A_{Ai} denotes the piston-side area of the cylinder and A_{Bi} is the rod-side area. Minimising this cost gives the flexibility to maximally tap into the returning flow from the cylinder. For example, as cylinder one extends, the returning flow from the meter-out side can be used for simultaneously extending actuators. Making full use of this in real-time would require a hardware overhaul in conventional systems and possibly a type of model predictive control. However, with regard to the goal of automating some construction crane work cycles, the globally optimal joint trajectories could be computed beforehand with the offline approach presented using Eq. (37) as the cost function.

Although we focused on traditional hydraulic systems, the proposed hydraulic energy minimisation strategy has practical relevance in some other hydraulic systems. For example, the flow rate delivered by the dedicated pumps in valveless systems may be minimised using the analysed approach to decrease the pumping effort. In addition, the results naturally extend to some hydraulic transformer or multi-chamber cylinder systems operating on a single constant pressure source. For example, these multi-chamber systems' energy efficiency is better compared to traditional systems because the displaced cylinder area can be discretely controlled to reduce the pressure losses over the control valve. Using the efficient flow-reducing strategy, the inputted hydraulic energy of the multi-chamber system can be minimised effectively because manipulator dynamics are not required. However, the minimised flow objective is complicated in this case by the variety of choices for the cylinder areas in the multi-chamber cylinders. Overall, these cases imply that the flow-reducing optimal control presented should be practically relevant in future applications. Finally, the extension of the DP approach to non-planar manipulators with a redundant extension joint and a base-rotating actuator is possible in future work.

8. Conclusions

Because conventional methods can be insufficient when it comes to resolving kinematic redundancy at the hydraulic level, we proposed cost functions formulated into an effective, globally optimal approach to resolve the redundancy problem at the hydraulic level. The popular load-sensing and constant-pressure system architectures were treated. Due to the complex problem formulation required at the hydraulic level, the solutions obtained using particularly the point-wise optimal methods yielded significantly poorer results compared with the global approach, up to 15–30% greater energy use as seen in numerical examples. Furthermore, we found that a bootstrapped pseudospectral solution may produce practically comparable results to dynamic programming in terms of hydraulic energy minimisation. The results obtained are relevant to the typical planar feedback-controlled hydraulic construction crane, whose redundancy is founded on the prismatic reach function. However, extending the results to typical non-planar cranes is possible. Furthermore, we assumed a separate meter-in and separate meter-out valve configuration, which allowed negligible cylinder back-pressures in the closed-loop simulations presented. A differential connection to the hydraulic cylinder was not considered and we also assumed a fixed pump efficiency in the simulations. Considering a variable hydraulic pump efficiency could have influenced the simulation results and changed the outcome of this study. A variable pump efficiency, regular valve configuration and differential connection are to be considered in future work. Overall, the results obtained motivate a need and provide an efficient way to optimise the joint trajectories on prospective applications in which frequently driven work cycles for hydraulic construction cranes are automated.

Acknowledgement

This work was partially funded by the Graduate School of Concurrent Mechanical Engineering of Finland, Decision No. 141343 of the Academy of Finland. This funding is highly appreciated. The authors also wish to sincerely thank the anonymous reviewers for critically reading the manuscript and suggesting significant improvements.

Appendix A. Parameters

Table A.1
Workspace and dynamics parameters of the manipulator.

Parameter	Value
L_1	1.60 [m]
L_2	1.562 [m]
o_x	−0.225 [m]
o_y	0.957 [m]
r_1	0.771 [m]
r_2	0.663 [m]
r_3	1.856 [m]
m_1	80.11 [kg]
m_2	33.93 [kg]
m_3	570.19 [kg]
$q_{1_{min}} : q_{1_{max}}$	−0.212:1.517 [rad]
$q_{2_{min}} : q_{2_{max}}$	−2.545:−0.321 [rad]
$q_{3_{min}} : q_{3_{max}}$	0:1.04 [m]
$v_{i_{max}} : \forall i \in \{1,2,3\}$	0.2 [m/s]
$v_{i_{min}} : \forall i \in \{1,2,3\}$	−0.2 [m/s]
$a_{i_{max}} : \forall i \in \{1,2,3\}$	0.5 [m/s^2]
$a_{i_{min}} : \forall i \in \{1,2,3\}$	−0.5 [m/s^2]

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.autcon.2016.09.006.

References

[1] J. Koivumäki, J. Mattila, High performance nonlinear motion/force controller design for redundant hydraulic construction crane automation, Autom. Constr. 51 (2015) 59–77. http://dx.doi.org/10.1016/j.autcon.2014.12.014.
[2] M. Vukobratovic, M. Kircanski, A dynamic approach to nominal trajectory synthesis for redundant manipulators, IEEE Trans. Syst. Man Cybern. SMC-14 (4) (1984) 580–586. http://dx.doi.org/10.1109/TSMC.1984.6313329.
[3] B. Löfgren, Kinematic Control of Redundant Knuckle Booms With Automatic Path-following Functions, Royal Institute of Technology, KTH, Sweden, 2009, Ph.D. thesis. http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Akth%3Adiva-11495. 978-91-7415-513-6.
[4] P. Hołobut, Time-optimal control of hydraulic manipulators with path constraints, J. Theor. Appl. Mech. 43 (3) (2005) 523–538. ISSN 1429-2955.
[5] Y. Halevi, E. Carpanzano, G. Montalbano, Minimum energy control of redundant linear manipulators, J. Dyn. Syst. Meas. Control. 136 (2014) 051016–1. http://dx.doi.org/10.1115/1.4027419.
[6] A. Deo, I. Walker, Minimum effort inverse kinematics for redundant manipulators, IEEE Trans. Robot. Autom. 13 (5) (1997) 767–775. http://dx.doi.org/10.1109/70.631238.
[7] L. Sciavicco, B. Siciliano, Modelling and Control of Robot Manipulators, Springer. 2000. http://dx.doi.org/10.1007/978-1-4471-0449-0. 978-1-85233-221-1.
[8] L. Beiner, J. Mattila, An improved pseudoinverse solution for redundant hydraulic manipulators, Robotica 17 (2) (1999) 173–179. http://dx.doi.org/10.1017/S0263574799001216.
[9] G. Zeiger, A. Akers, Dynamic analysis of an axial piston pump swashplate control, Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci. 200 (1) (1986) 49–58. http://dx.doi.org/10.1243/PIME_PROC_1986_200_093_02.
[10] N.D. Manring, F.A. Damtew, The control torque on the swash plate of an axial-piston pump utilizing piston-bore springs, J. Dyn. Syst. Meas. Control. 123 (3) (2001) 471–478. http://dx.doi.org/10.1115/1.1386654.
[11] P. Achten, Dynamic high-frequency behaviour of the swash plate in a variable displacement axial piston pump, Proc. Proc. Inst. Mech. Eng., I: J. Syst. Control Eng. 227 (6) (2013) 529–540. http://dx.doi.org/10.1177/0959651813483419.
[12] P. Krus, On Load Sensing Fluid Power Systems: With Special Reference to Dynamic Properties and Control Aspects, PhD Thesis, Linköping University, Sweden, 1988. 9178704022.
[13] H.C. Pedersen, T.O. Andersen, M.R. Hansen, Controlling a conventional LS-pump based on electrically measured LS-pressure, Proceedings of the Bath/ASME Symposium on Fluid Power and Motion Control, 2008, pp. 531–546.
[14] M. Axin, B. Eriksson, P. Krus, Flow versus pressure control of pumps in mobile hydraulic systems, Proc. Inst. Mech. Eng., I: J. Syst. Control Eng. 228 (4) (2014) 245–256. http://dx.doi.org/10.1177/0959651813512820.
[15] J. Koivumäki, J. Mattila, Stable and high performance energy-efficient motion control of electric load sensing controlled hydraulic manipulators, Proceedings of the ASME/BATH Symposium on Fluid Power and Motion Control, American Society of Mechanical Engineers, 2013, pp. V001T01A024-V001T01A024. http://dx.doi.org/10.1115/FPMC2013-4440.
[16] O. Sundstrom, L. Guzzella, A generic dynamic programming Matlab function, Proceedings of the IEEE Control Applications (CCA) & Intelligent Control (ISIC), 2009, pp. 1625–1630. http://dx.doi.org/10.1109/CCA.2009.5281131.
[17] D.E. Kirk, Optimal Control Theory: An Introduction, Courier Dover Publications. 2012. 0486434842.
[18] W. Shen, J. Jiang, X. Su, H.R. Karimi, Parameter matching analysis of hydraulic hybrid excavators based on dynamic programming algorithm, J. Appl. Math. 2013 (2013) http://dx.doi.org/10.1155/2013/615608.
[19] P. Elbert, S. Ebbesen, L. Guzzella, Implementation of dynamic programming for n-dimensional optimal control problems with final state constraints, IEEE Trans. Control Syst. Technol. 21 (3) (2013) 924–931. http://dx.doi.org/10.1109/TCST.2012.2190935.
[20] M.H. Choi, Redundancy resolution by minimization of joint disturbance torque for independent joint controlled manipulators, Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), 1999, pp. 392–397. http://dx.doi.org/10.1109/AIM.1999.803199.
[21] C. Pholsiri, Task-based Decision Making and Control of Robotic Manipulators, University of Texas, USA, 2004. Ph.D. thesis
[22] S. Aranovskiy, A. Losenkov, C. Vazquez, Position control of an industrial hydraulic system with a pressure compensator, Proceedings of the 22nd Mediterranean Conference of Control and Automation (MED), 2014, pp. 1329–1334. http://dx.doi.org/10.1109/MED.2014.6961560.
[23] R.N. Jazar, Theory of Applied Robotics: Kinematics, Dynamics, and Control, Springer, 2010. http://dx.doi.org/10.1007/978-1-4419-1750-8. 978-1-4419-1749-2.
[24] I.M. Ross, F. Fahroo, Pseudospectral knotting methods for solving nonsmooth optimal control problems, J. Guid. Control. Dyn. 27 (3) (2004) 397–405. http://dx.doi.org/10.2514/1.3426.
[25] I.M. Ross, M. Karpenko, A review of pseudospectral optimal control: from theory to flight, Annu. Rev. Control. 36 (2) (2012) 182–197. http://dx.doi.org/10.1016/j.arcontrol.2012.09.002.
[26] F. Flacco, A. De Luca, O. Khatib, Motion control of redundant robots under joint constraints: saturation in the null space, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2012, pp. 285–292. http://dx.doi.org/10.1109/ICRA.2012.6225376.
[27] S. Liu, B. Yao, Programmable valves: a solution to bypass deadband problem of electro-hydraulic systems, Proceedings of the American Control Conference, vol. 5, 2004, pp. 4438–4443.

[28] S. Liu, B. Yao, Coordinate control of energy saving programmable valves, IEEE Trans. Control Syst. Technol. 16 (1) (2008) 34–45. http://dx.doi.org/10.1109/TCST.2007.903073.

[29] Q. Yuan, J.Y. Lew, Modeling and control of two stage twin spool servo-valve for energy-saving, Proceedings of the American Control Conference, IEEE, 2005, pp. 4363–4368. http://dx.doi.org/10.1109/ACC.2005.1470666.

Publication VI

Jarmo Nurmi, Jouni Mattila, "Global energy-optimal redundancy resolution of hydraulic manipulators: Experimental results for a forestry manipulator." *Energies*, vol. 10, no. 5, 31 p., May 2017.

This is an Accepted Open-Access Article published by MDPI in Energies on 05/2017, available online: `http://dx.doi.org/10.3390/en10050647`

energies

MDPI

Article

Global Energy-Optimal Redundancy Resolution of Hydraulic Manipulators: Experimental Results for a Forestry Manipulator

Jarmo Nurmi * and Jouni Mattila

Laboratory of Automation and Hydraulic Engineering, Tampere University of Technology, P.O. Box 589 (Korkeakoulunkatu 6), FIN-33101 Tampere, Finland; jouni.mattila@tut.fi

* Correspondence: jarmo.nurmi@tut.fi

Academic Editor: Andrea Vacca

Received: 15 March 2017; Accepted: 3 May 2017; Published: 6 May 2017

Abstract: This paper addresses the energy-inefficiency problem of four-degrees-of-freedom (4-DOF) hydraulic manipulators through redundancy resolution in robotic closed-loop controlled applications. Because conventional methods typically are local and have poor performance for resolving redundancy with respect to minimum hydraulic energy consumption, global energy-optimal redundancy resolution is proposed at the valve-controlled actuator and hydraulic power system interaction level. The energy consumption of the widely popular valve-controlled load-sensing (LS) and constant-pressure (CP) systems is effectively minimised through cost functions formulated in a discrete-time dynamic programming (DP) approach with minimum state representation. A prescribed end-effector path and important actuator constraints at the position, velocity and acceleration levels are also satisfied in the solution. Extensive field experiments performed on a forestry hydraulic manipulator demonstrate the performance of the proposed solution. Approximately 15–30% greater hydraulic energy consumption was observed with the conventional methods in the LS and CP systems. These results encourage energy-optimal redundancy resolution in future robotic applications of hydraulic manipulators.

Keywords: hydraulic manipulator; redundancy resolution; energy optimisation; energy-optimal; global; dynamic programming; actuator limits; load-sensing system; constant-pressure system; mobile hydraulic valve

1. Introduction

Hydraulic loaders are powerful, typically 4-degrees-of-freedom (4-DOF) manipulators, that are used for heavy-duty material handling on mobile machines, such as trucks and forest forwarders. Currently, hydraulic loaders are predominantly operated by humans, but loader manufacturers are interested in diversifying their offerings through robotic functionality to remain competitive. On the research side, and in the industry, subjects dealing with reducing reliance on the expertise of human operators in loader control have consequently received much attention. Several next-generation robotic semi-automated approaches, for example, have been proposed to help alleviate the operator's workload by automating the repetitive work tasks that would normally be the operator's responsibility [1,2]. On one hand, these semi-automated approaches discuss closed-loop control, and on the other hand, they involve redundancy resolution because hydraulic loaders are typically kinematically redundant. Although both problems are important, it is safe to estimate that redundancy resolution of hydraulic manipulators has received less and too little interest in academia in relation to the importance of redundancy resolution compared to closed-loop control problems. This is probably because so many different types of redundancy resolutions [3–9] have been proposed for general

systems that the general understanding is that these redundancy resolutions can effectively reduce the energy consumption of hydraulic manipulators, for example, through the minimum velocity norm or weighted pseudo-inverse solutions [10]. One of the aims in this paper is to show that these general redundancy resolutions can be sub-optimal at the valve-controlled hydraulic system level.

Moreover, because there are many different redundancy resolution strategies for different objectives, choosing the most compelling objective for a hydraulic manipulator is not always straightforward. Looking at the handful of redundancy resolutions dealing with specific problems and objectives pertaining to hydraulic manipulators, however, it becomes evident that global energy-optimal redundancy resolutions are lacking. To explain the importance of global energy-optimal redundancy resolution, system energy inefficiency is a major problem in hydraulic manipulators [11]. Therefore, to address this problem, we resolve the redundancy in a highly desirable hydraulic energy-optimal manner and effectively treat the inherently limited joint motions of hydraulic manipulators by enforcing important actuator position, velocity and acceleration limits in our solution.

1.1. Literature Review

In the literature, Mettin et al. [12], Löfgren [13] and Ortiz Morales et al. [1] addressed the joint-limited minimum-time redundancy resolution of hydraulic manipulators, and Löfgren [13] has also studied redundancy resolution which maximises load capacity. Furthermore, Beiner and Mattila [10] have derived a pseudo-inverse method for redundancy resolution in the hydraulic manipulator's actuator space to minimise the cylinder velocity norm or the weighted kinetic energy instantaneously. In addition, Flacco et al. [9] have proposed a joint-limited pseudo-inverse-based algorithm, which is also suitable for hydraulic applications, that saturates joint velocities in the null space. These different solutions, however, do not discuss hydraulic energy-optimality. In addition, generally minimising the energy consumption of actuators, as proposed by Vukobratovic and Kircanski [14] as a local solution and developed by Hirakawa and Kawamura [15] as an optimal control problem, can be a sub-optimal strategy for hydraulic manipulators. This is a direct consequence of the fact that the actuators' energy consumption does not equate to the energy consumption of the typical valve-controlled load-sensing (LS) and constant-pressure (CP) hydraulic systems because the pressure losses encountered in these systems are not considered. It is important to derive energy-optimal redundancy resolutions for these commonly used systems. Whereas LS systems are predominantly used in hydraulic commercial manipulators in general, CP systems are used mostly in robotic closed-loop controlled applications and next-generation energy-efficient secondary-controlled applications because of the systems' superior damping and simplicity compared with the LS systems [16–18].

1.2. Paper Contribution and Organization

In this paper, we minimise the energy consumption of the widely used LS and CP hydraulic systems through redundancy resolution to address the major problem of energy-inefficiency in hydraulic manipulators. This solution is in stark contrast to known conventional methods that are not energy-optimal at the valve-controlled actuator and hydraulic power system interaction level. To summarise, our redundancy resolution entails moving the manipulator hydraulic cylinders in an energy-optimal manner while the manipulator's end-effector follows a prescribed path, and actuator limits at the position, velocity and acceleration levels are satisfied. To address both of these hydraulic systems optimally, we propose a tailored cost function of the energy consumption for each. For example, the simpler CP system is shown to require minimisation of the pump flow rate. This objective leads to our effective solution in which the cylinder's effective areas scale the cylinder velocities in a discontinuous manner. Minimising the CP system energy cost, hence, intuitively amounts to minimising the cylinder volumes displaced in the solution. The cylinder areas among the manipulator's hydraulic cylinders, including the different piston and rod-side areas of the asymmetric cylinders, render minimisation of the CP system's energy consumption, in particular, a highly tractable objective.

With the objectives in place, the redundancy could obviously be resolved by formulating the optimal control problem so that the motions of all the joints are optimised subject to specific motion constraints. However, that approach would result in a high-dimensional problem, which would be unsuitable for a global dynamic programming (DP) solution and generally inconvenient. To reduce the problem dimensionality and computational complexity, in this study, only the motion of the redundant extension cylinder is directly optimised. The optimal non-redundant joint motions are, instead, resolved by employing non-singular inverse kinematics, motivated by the formulation used by Choi [19] and Flacco et al. [20]. Moreover, in the present study, hydraulic system properties are incorporated in the cost functions proposed in steady-state equations. This approach results in our desirable minimum-state problem, which requires only two states and one control and which uses the computational resources available much more sparingly than the approach of simultaneously optimising all of the joint motions. However, because of the modelling simplifications in our problem formulation, a natural concern that could arise is whether significant hydraulic aspects of the problem are neglected. Thus, extensive experiments, which compare the energy consumptions of the redundancy resolutions proposed in this study against other global and instantaneous strategies in the LS and CP systems is presented.

This energy-optimal solution is primarily suitable as an offline approach, much like the minimum-time solutions that were proposed by Ortiz Morales et al. [1] and Mettin et al. [12], and it is practical as such. Extensive measurements by Ortiz Morales et al. [1], Mettin et al. [12] and Westerberg [2] have shown that the typical workspace activity for 4-DOF manipulator is concentrated on specific areas, which consequently means that there are in practice a limited number of significant point-to-point movements to optimise with respect to energy consumption. However, at the implementation stage, it might be feasible to pre-compute the energy-optimal trajectories over a few time horizons for all of these significant point-to-point movements so that it would be possible to choose between slower and faster energy-optimal trajectories in real-time. Furthermore, with our proposed methodology, one significant implication is that only the redundant joint's trajectories are necessary for real-time applications. If having more storage space were considered more important than having more computational power reserve, then the non-redundant motion trajectories could be resolved online from the remaining non-singular inverse kinematics.

This paper is an extension of our previous work [21]. As an extension of previous work, redundancy resolution is expanded from vertical plane three-degrees-of-freedom (3-DOF) serial-link hydraulic manipulators to complex 4-DOF serial-link hydraulic manipulators, which include boom base motion. The system model proposed for the energy-optimal solution is also systematically presented in this study. A third contribution of this paper compared to our previous simulation study concerning 3-DOF serial-link hydraulic manipulators is the extensive field experimental evaluation of the redundancy resolution proposed for 4-DOF serial-link hydraulic manipulators and its comparison to existing conventional methods. In this experimental evaluation, the performance of the proposed DP solution is compared to a "bootstrapped" DIDO solution [22] by employing the proposed problem formulation. The performance of the globally hydraulic energy-optimised redundancy resolution has not been previously experimentally evaluated.

This paper is organized as follows. First, the test system, which is a commercial forestry hydraulic manipulator, is described in Section 2. Second, a simplified mathematical model of this complex manipulator is presented in Section 3. In Section 4, the hydraulic manipulator's optimal control problem is reduced to the bare essentials with the minimum-state formulation. The problem's desirable global dynamic programming solution is discussed in Section 5. Finally, we experimentally evaluate the proposed redundancy resolution by comparing its performance to that of conventional methods in Section 6. The experimental results are from the real test-bed introduced in Section 2. Conclusions are provided in Section 7.

2. Test-Bed

Our test-bed is the K70M hydraulic manipulator on the Ponsse Caribou S10 8-wheel forest forwarder (Figure 1). In the following, we describe this manipulator in more detail.

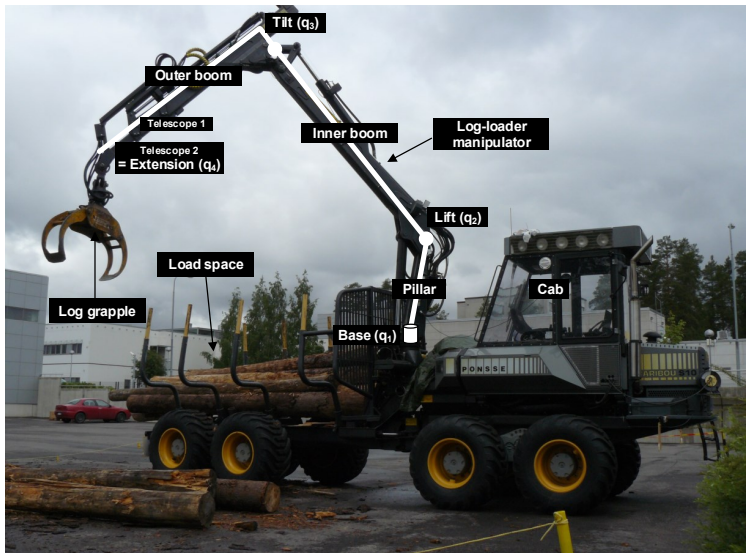


Figure 1. Kinematically redundant K70M hydraulic manipulator on Ponsse Caribou S10 forwarder, with log grapple end-effector.

2.1. System Description

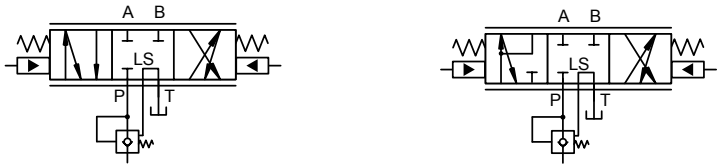
The K70M hydraulic manipulator is a standard log-loader that has a grapple end-effector attached to the crane tip. The manipulator has four principal joints, which consist of the base, lift, tilt and extension joints shown in Figure 1. These joints are actuated by using proportional valve-controlled hydraulic cylinders (see Table 1). The manipulator also has two telescoping links in series to extend its reachability. Measuring from the base to the tip, the manipulator can extend to approximately 9.1 m, and at this distance, its load-carrying capacity is 650 kg.

Table 1. Hydraulic cylinder dimensions of the test-bed manipulator.

i	$\oslash A_{\mathrm{A}i}/A_{\mathrm{B}i}$ − Stroke [mm]	Explanation
1	$2 \times \oslash 115/115 - 720$	Base cylinders
2	$\oslash 120/70 - 610$	Lift cylinder
3	$\oslash 115/60 - 725$	Tilt cylinder
4	$\oslash 70/45 - 1850$	Extension cylinder

A two-stage, pressure-compensated Parker K170LS mobile hydraulic valve is used to control the manipulator's cylinders; see Figure 2, where the valve's spool types are illustrated. The mobile hydraulic valve also includes very complex in-built hydro-mechanical functionality, such as a standard counter pressure valve, which raises the tank pressure, and anti-cavitation valves in each actuator port (not shown in Figure 2). In the over-running load case, in particular, these anti-cavitation valves allow any fluid flow rate from the tank line to the non-loaded cylinder chamber to exist before the chamber pressure drops to a level below a preset minimum hydraulic cylinder pressure. This hydro-mechanical functionality can be significant since these anti-cavitation valves may reduce pump flow rate requirement.

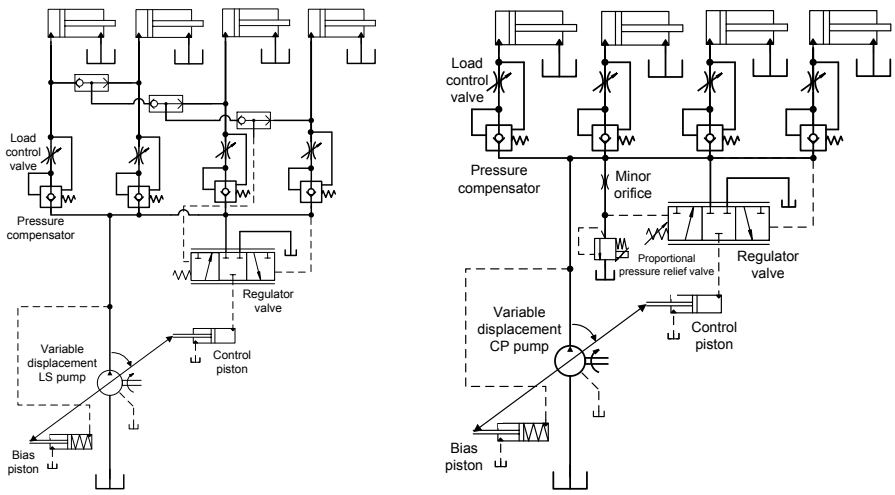
The manipulator, as an industrial standard, is powered by a LS hydraulic pump; see [23–25]. Therefore, an electro-hydraulic CP system was built in parallel to allow comparison to the LS system for the experiments. Hence, the same hydraulic pump and pump LS regulator valve can be used to control the pump displacement. The hydraulic pump is an A11VLO130DRS variable axial piston pump from Bosch Rexroth with a maximum flow rate of 325 dm^3/min at 2500 rotations per minute (including the charge pump flow) and 21.5 MPa maximum supply pressure.



(**a**) (**b**)

Figure 2. Mobile hydraulic spool valves of the test-bed manipulator: (**a**) Typical mobile hydraulic spool, used with the base, lift and tilt cylinders; and (**b**) differential spool (see position on the left) used with the extension cylinder.

The LS system's conventional low-level operating principle, which is reported in [23], is illustrated in Figure 3a. In this system, the maximum driven chamber pressure among the base, lift, tilt and extension cylinders is fed back to the variable displacement pump's hydro-mechanical controller, which sets the pump supply pressure a LS pressure margin Δp_{LS} above this feedback value. With the CP electro-hydraulic control method of the hydro-mechanical feedback system shown in Figure 3b, on the other hand, the pressure reference for the variable displacement pump's hydro-mechanical controller is generated directly from the supply pressure based on the electrical input of the proportional pressure relief valve between the minor (0.7 mm) orifice and the tank-connected pressure valve. This control method is called "remote control" in the Bosch Rexroth catalog [26]. We have simplified the symbol of the metre-in side load control valve and omitted the metre-out side for clarity.



(**a**) (**b**)

Figure 3. Hydraulic systems of the test-bed manipulator: (**a**) Conventional load-sensing system; and (**b**) electro-hydraulic constant-pressure system.

2.2. Instrumentation and Data Acquisition

The hydraulic manipulator was equipped with multiple sensors for the experiments:

(1) the base angle was measured using an RCC absolute angle resolver from LTN with a resolution of 0.01°;
(2) the lift and tilt angles were measured using incremental angle encoders ROD 456–5000 from Heidenhain with resolutions of 0.0072° and 0.0029°, after 10-fold and 25-fold interpolation;
(3) the extension cylinder position was measured with a DG 60 L incremental position encoder from Stegmann with a resolution of 160 μm;
(4) the pump supply pressure and, for identification purposes, cylinder pressures were measured using an NAH250 analogue pressure transducer from Trafag with an accuracy of ±0.1% from full-scale 25 MPa and
(5) the pump flow rate was measured using a gear type flow meter VC 5 F2 PS from Kracht with a resolution of 0.005 dm^3 and a range of −250 to 250 dm^3 per minute.

The joint encoder installations are shown in Figure 4. These sensor measurements were recorded at a 2 ms sampling rate, and the hydraulic valves were controlled via CAN using a dSPACE MicroAutoBox II processing unit for the experimental evaluation.

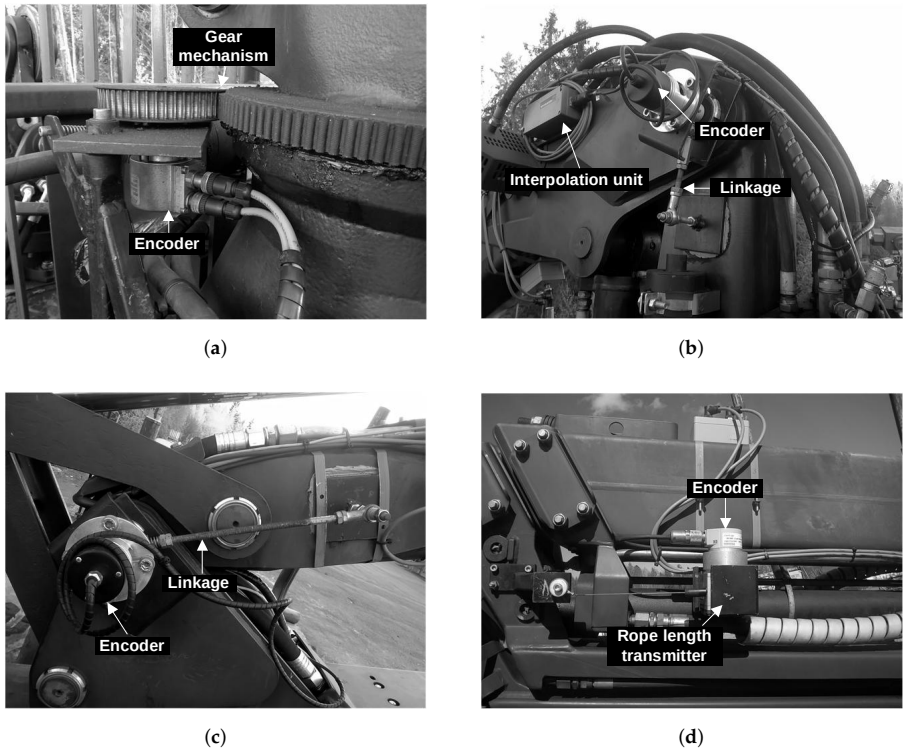


Figure 4. Joint encoders installed at the principal joints of the hydraulic manipulator: (**a**) Base angle (q_1) encoder; (**b**) lift angle (q_2) encoder; (**c**) tilt angle (q_3) encoder; and (**d**) extension length (q_4) encoder.

3. Model of the Kinematically Redundant Four-Degrees-of-Freedom Hydraulic Manipulator

Next, the test-bed is described in more detail. For redundancy resolution of this test-bed manipulator, the mathematical model of the manipulator is systematically presented in the following, with particular attention paid to the applicability of the model for optimising energy consumption.

3.1. Forward Kinematics in the Joint Space

The hydraulic manipulator is shown in Figure 5. The joint-position vector of this manipulator can be defined as follows:

$$\mathbf{q} = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix}^{\mathrm{T}} \tag{1}$$

where q_1 denotes the base angle, q_2 denotes the lift angle, q_3 denotes the tilt angle and q_4 is the total extension length (a prismatic joint). These are the principal hydraulically actuated joints. All q_i are taken positive counter-clockwise.

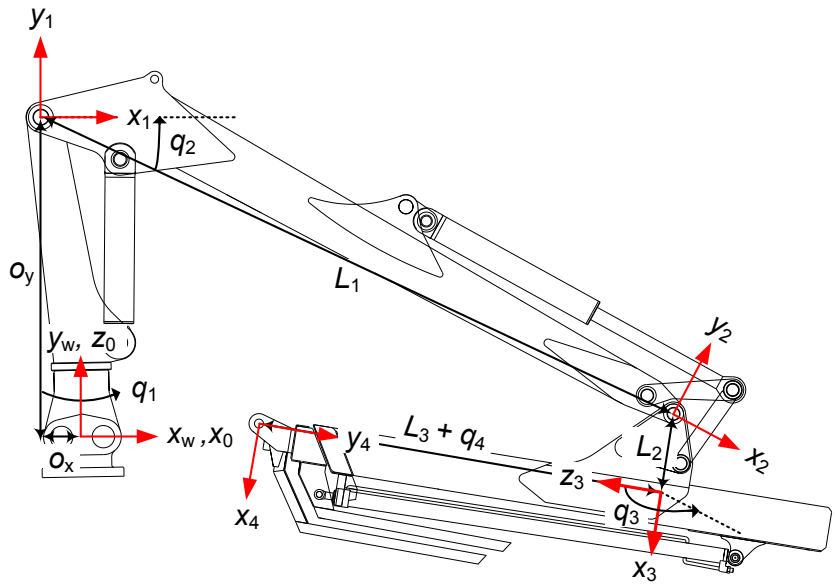


Figure 5. 4-DOF kinematically redundant hydraulic manipulator, with coordinate frames attached to the links based on the classical Denavit-Hartenberg convention.

The manipulator's end-effector position can be expressed from the base frame 0 to frame 4 in accordance with the classical Denavit-Hartenberg (DH) convention [27]. By using the DH transformation matrix $\mathbf{A}_i^{i-1}$, which determines the coordinate transformation between successive frames, we get:

$$\mathbf{A}_4^0 = \mathbf{A}_1^0 \mathbf{A}_2^1 \mathbf{A}_3^2 \mathbf{A}_4^3 \tag{2}$$

as the overall coordinate transformation expressed in the base frame. The transformation matrix $\mathbf{A}_i^{i-1} \in \mathbb{R}^{4x4}$ used consecutively is defined as:

$$\mathbf{A}_i^{i-1} = \begin{bmatrix} \mathrm{c}_{\theta_i} & -\mathrm{s}_{\theta_i}\mathrm{c}_{\alpha_i} & \mathrm{s}_{\theta_i}\mathrm{s}_{\alpha_i} & a_i\mathrm{c}_{\theta_i} \\ \mathrm{s}_{\theta_i} & \mathrm{c}_{\theta_i}\mathrm{c}_{\alpha_i} & -\mathrm{c}_{\theta_i}\mathrm{s}_{\alpha_i} & a_i\mathrm{s}_{\theta_i} \\ 0 & \mathrm{s}_{\alpha_i} & \mathrm{c}_{\alpha_i} & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where, for example, c_{α_i} denotes $\cos(\alpha_i)$ and s_{α_i} denotes $\sin(\alpha_i)$. The DH parameters given in Table 2 determine these matrix elements.

By transforming from the base frame to the x_w, y_w and z_w coordinates, which denote the x-, y- and z-positions of the end-effector expressed in the global frame, respectively, we get:

$$x_{\mathrm{w}} = \cos(q_1)\{L_1\cos(q_2) - L_2\sin(q_2+q_3) + (L_3+q_4)\cos(q_2+q_3) + o_{\mathrm{x}}\} \quad (4)$$

$$y_{\mathrm{w}} = L_1\sin(q_2) + L_2\cos(q_2+q_3) + (L_3+q_4)\sin(q_2+q_3) + o_{\mathrm{y}} \text{ and} \quad (5)$$

$$z_{\mathrm{w}} = \sin(q_1)\{L_1\cos(q_2) - L_2\sin(q_2+q_3) + (L_3+q_4)\cos(q_2+q_3) + o_{\mathrm{x}}\} \quad (6)$$

as the final end-effector position, which we denote with $\mathbf{x}_{\mathrm{t}} = [x_{\mathrm{w}}\ y_{\mathrm{w}}\ z_{\mathrm{w}}]^{\mathrm{T}}$. Constants L_1 and L_3 denote the manipulator link lengths, L_2 is an offset between the tilt joint and the extension link and o_{x} and o_{y} are the offsets in the direction of the x- and y-axes, respectively, from the global frame to the lift joint. The workspace of the test-bed manipulator in the vertical plane is depicted in Figure 6 by using the parameters given in Table 2.

Table 2. Denavit-Hartenberg parameters of the 4-DOF hydraulic manipulator.

Frame No.	a_i [m]	α_i [rad]	d_i [m]	θ_i [rad]
1	$o_{\mathrm{x}} = -0.208$	$\pi/2$	$o_{\mathrm{y}} = 1.605$	q_1
2	$L_1 = 3.504$	0	0	q_2
3	$L_2 = 0.376$	$\pi/2$	0	$q_3 + \pi/2$
4	0	$-\pi/2$	$L_3 + q_4 = 2.13 + q_4$	0

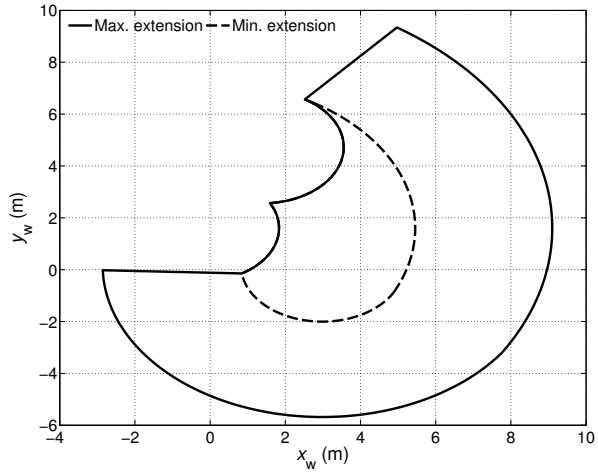


Figure 6. Theoretical reachable workspaces of the K70M hydraulic manipulator in this vertical plane. Reachable workspaces with maximum ($q_4 = 3.7$ m) and minimum extension ($q_4 = 0$ m) have been presented.

Differentiating $\mathbf{x}_{\mathrm{t}}$ in view of Equations (4)–(6) with respect to time yields the end-effector velocity:

$$\dot{\mathbf{x}}_{\mathrm{t}} = \mathbf{J}(\mathbf{q})\dot{\mathbf{q}} = \mathbf{J}_1\dot{q}_1 + \mathbf{J}_2\dot{q}_2 + \mathbf{J}_3\dot{q}_3 + \mathbf{J}_4\dot{q}_4 \quad (7)$$

where $\dot{\mathbf{x}}_{\mathrm{t}} = [\dot{x}_{\mathrm{w}}\ \dot{y}_{\mathrm{w}}\ \dot{z}_{\mathrm{w}}]^{\mathrm{T}}$, the joint velocity vector $\dot{\mathbf{q}} = [\dot{q}_1\ \dot{q}_2\ \dot{q}_3\ \dot{q}_4]^{\mathrm{T}}$ and the Jacobian matrix $\mathbf{J}(\mathbf{q})$ is the partial derivative of $\mathbf{x}_{\mathrm{t}}$ with respect to $\mathbf{q}$. This equation was also expressed by using the column vectors $\mathbf{J}_i$, for $i \in \{1,2,3,4\}$, for future reference. Differentiating $\dot{\mathbf{x}}_{\mathrm{t}}$ once more yields:

$$\ddot{\mathbf{x}}_{\mathrm{t}} = \dot{\mathbf{J}}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{J}(\mathbf{q})\ddot{\mathbf{q}} \quad (8)$$

with the end-effector acceleration $\ddot{\mathbf{x}}_{\mathrm{t}}$ and the joint acceleration vector $\ddot{\mathbf{q}}$ defined similarly to Equation (7). The elements of the Jacobian time derivative $\dot{\mathbf{J}}$ can be resolved from:

$$\dot{J}(\mathbf{q},\dot{\mathbf{q}})_{ji} = \frac{\partial J(\mathbf{q})_{ji}}{\partial \mathbf{q}}\dot{\mathbf{q}} = \sum_{n=1}^{4} \frac{\partial J(\mathbf{q})_{ji}}{\partial q_n}\dot{q}_n \tag{9}$$

over the workspace dimensions $j \in \{1,2,3\}$ and the joint dimensions $i \in \{1,2,3,4\}$. Because the Jacobian matrix $\mathbf{J}(\mathbf{q}) \in \mathbb{R}^{3\times 4}$ is non-square, the inverse kinematics clearly have an infinite number of solutions inside the workspace. Thus, a redundancy resolution is required.

3.2. Joint Space to Actuator Space Transformations

Transformations, which yield the cylinder coordinates as a function of the joints, are presented next. By using the chain rule, we obtain:

$$c_i = c_i(q_i), \tag{10}$$

$$v_i = \frac{\partial c_i(q_i)}{\partial q_i}\dot{q}_i = r_i\dot{q}_i \text{ and} \tag{11}$$

$$a_i = \frac{\partial r_i}{\partial q_i}\dot{q}_i^2 + r_i\ddot{q}_i \tag{12}$$

as the general equations, where $c_i(q_i)$, for the actuator index $i \in \{1,2,3,4\}$, is the cylinder length, v_i is the cylinder (piston) velocity, a_i is the cylinder (piston) acceleration, not to be confused with the DH parameter, and r_i is the cylinder torque arm. The joint space to actuator space transformations of the test-bed are presented in the following.

Firstly, the base cylinder has a constant torque arm r_1 because a rack-and-pinion mechanism, as illustrated in Figure 7, is used (Equation (11)). Furthermore, the torque arm r_2 of the lift cylinder can be written as:

$$\begin{aligned} \beta_{13} &= q_2 - \beta_{11} - \beta_{12} + \pi/2 \\ r_2(q_2) &= \frac{L_{11}L_{12}\sin(\beta_{13})}{c_2} = \frac{L_{11}L_{12}\sin(\beta_{13})}{\sqrt{L_{11}^2 + L_{12}^2 - 2L_{11}L_{12}\cos(\beta_{13})}} \end{aligned} \tag{13}$$

as a function of the lift angle q_2 and the lift cylinder length c_2. The constants L_{11}, L_{12}, β_{11} and β_{12} and the joint angle q_2 are illustrated in Figure 8a.

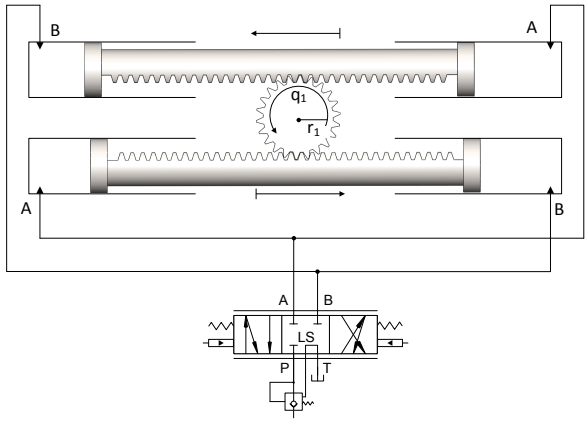


Figure 7. Rack-and-pinion mechanism at the test-bed manipulator's base controlled using pressure-compensated proportional valve.

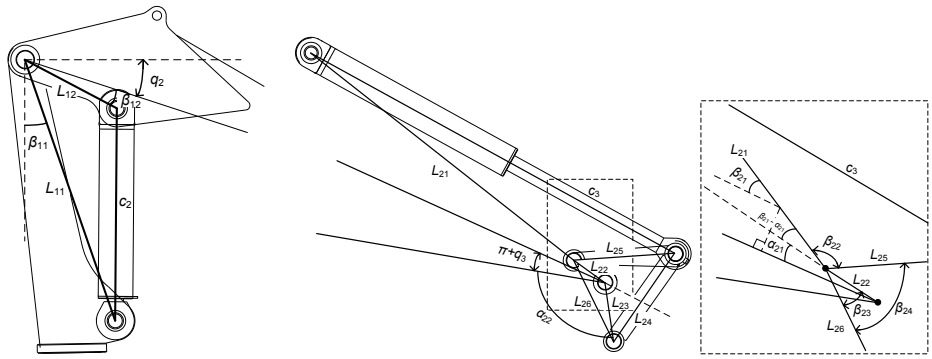


Figure 8. Joint actuation mechanisms of the test-bed manipulator: (**a**) Lift joint and associated variables; and (**b**) tilt joint and associated variables, with a close-up.

The four-bar linkage of the tilt cylinder is governed by more complex kinematic equations. These equations, from the tilt angle q_3 to tilt cylinder length c_3, can be written as follows:

$$\begin{aligned}
\alpha_{23} &= \pi + q_3 + \alpha_{21} + \alpha_{22} \\
L_{26} &= \sqrt{L_{22}^2 + L_{23}^2 - 2L_{22}L_{23}\cos(\alpha_{23})} \\
\beta_{23} &= \operatorname{asin}\big(L_{23}/L_{26}\sin(\alpha_{23})\big) \\
\beta_{24} &= \operatorname{acos}\big((L_{25}^2 + L_{26}^2 - L_{24}^2)/(2L_{25}L_{26})\big) \\
\beta_{22} &= \pi - \beta_{21} - \beta_{24} + \beta_{23} + \alpha_{21} \\
c_3(q_3) &= \sqrt{L_{21}^2 + L_{25}^2 - 2L_{21}L_{25}\cos(\beta_{22})} \\
&\approx s_5 q_3^5 + s_4 q_3^4 + s_3 q_3^3 + s_2 q_3^2 + s_1 q_3 + s_0 \\
&= f_q(\mathbf{s}, q_3),
\end{aligned} \tag{14}$$

where α_{21}, α_{22}, β_{21}, L_{21}, L_{22}, L_{23}, L_{24} and L_{25} are the constants and β_{22}, β_{23}, β_{24} and L_{26} are the variables (see Figure 8b, where q_3 is negative). Laws of cosines and sines were used. A quintic function $f_q(\mathbf{s}, q_3)$, with $\mathbf{s} \in \mathbb{R}^6$ as the coefficient vector, can be fitted to the c_3 data to simplify the equation of the tilt torque arm r_3. Hence, differentiating c_3 with respect to time in view of Equation (11) yields:

$$\begin{aligned}
v_3 &= -\frac{\partial f_q(\mathbf{s}, q_3)}{\partial q_3}\dot{q}_3 \\
&= -(5s_5 q_3^4 + 4s_4 q_3^3 + 3s_3 q_3^2 + 2s_2 q_3 + s_1)\dot{q}_3 \\
&= -r_3(q_3)\dot{q}_3
\end{aligned} \tag{15}$$

as the tilt cylinder's piston velocity. Multiplication by minus one considers the opposite signs of the piston and angular velocities. The "torque arm" r_4 of the telescoping extension mechanism is 1/2 because the extension cylinder velocity is half of the joint velocity due to the telescope's mechanical multiplier.

3.3. Cylinder Forces

Manipulator dynamics are considered through the hydraulic cylinder forces, which can be resolved from the inverse rigid-body dynamics as follows:

$$\mathbf{F} = \mathbf{R}^{-1}\big(\underbrace{\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) + \mathbf{F}_{\mathrm{f}}(\mathbf{q}, \dot{\mathbf{q}})}_{\boldsymbol{\tau}}\big) \tag{16}$$

where the cylinder force vector $\mathbf{F}$ equals $[F_1\ F_2\ F_3\ F_4]^{\mathrm{T}}$, $\boldsymbol{\tau}$ is the corresponding cylinder torque vector, $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{4x4}$ denotes the positive definite, symmetric inertia matrix at joint position $\mathbf{q}$, $\mathbf{G}(\mathbf{q}) \in \mathbb{R}^4$ denotes the gravitational torque vector and $\mathbf{F}_{\mathrm{f}}(\mathbf{q}, \dot{\mathbf{q}}) \in \mathbb{R}^4$ includes the hydraulic cylinder and joint friction forces. The inertia matrix and the gravitational load vector can be resolved via well-known Lagrangian conventions by summing the dynamical contributions of each link in the global frame [27]. Resolving these components regarding our test-bed is mostly standard practice, with the exception that the extension cylinder's gravitational component must be doubled to consider the mechanical disadvantage of the telescope mechanism (The mechanism doubles the extension cylinder stroke). Because the centrifugal and Coriolis forces are in low magnitude, their effects can be omitted. Conventional exponential friction force models can be primarily exploited (see [28]), but the extension joint's friction force is shown to be accurately represented by using the developed unconventional model (see Appendix B for the test-bed's cylinder force modelling results).

The inverse torque arm matrix is defined as in:

$$\mathbf{R}^{-1} = \begin{bmatrix} 1/r_1 & 0 & 0 & 0 \\ 0 & 1/r_2 & 0 & 0 \\ 0 & 0 & 1/r_3 & 0 \\ 0 & 0 & 0 & 1/r_4 \end{bmatrix} \tag{17}$$

Consequently, these positive matrix elements, denoted by r_i^{-1}, are the inverses of the hydraulic cylinder torque arms. However, r_4^{-1} is a non-physical multiplier because the fourth joint is prismatic.

3.4. Steady-State Cylinder Pressures Using a Mobile Hydraulic Valve

Finally, moving on to the hydraulic system equations, we introduce simplifications to keep the system dimensionality at a manageable level for optimisation purposes. To introduce our subsequent simplifications, firstly, consider the pressure dynamics of the hydraulic cylinders denoted by [29]:

$$\dot{p}_{\mathrm{A}} = \frac{B_{\mathrm{eff}}}{V_{0\mathrm{A}} + x_{\mathrm{p}} A_{\mathrm{A}}}\left(Q_{\mathrm{A}} - v A_{\mathrm{A}}\right) \text{ and} \tag{18}$$

$$\dot{p}_{\mathrm{B}} = \frac{B_{\mathrm{eff}}}{V_{0\mathrm{B}} + (L_{\max} - x_{\mathrm{p}}) A_{\mathrm{B}}}\left(Q_{\mathrm{B}} + v A_{\mathrm{B}}\right) \tag{19}$$

where B_{eff} is the effective bulk modulus, $V_{0\mathrm{A}}$ and $V_{0\mathrm{B}}$ are the dead volumes, x_{p} is the piston position, v is the piston velocity, A_{A} is the piston-side area, A_{B} is the rod-side area and Q_{A} and Q_{B} are the flow rates to the cylinder chambers. Actuator indices were omitted for brevity. Solving the steady-state cylinder velocity from Equations (18) and (19) yields:

$$v_{\mathrm{ss}} = Q_{\mathrm{A}}/A_{\mathrm{A}} = -Q_{\mathrm{B}}/A_{\mathrm{B}} \tag{20}$$

Now, we omit these hydraulic system dynamics and focus on deriving the hydraulic system's steady-state pressure equations by utilising Equation (20). Because the flow rates to the hydraulic cylinders are controlled by using the mobile hydraulic valve's two vastly different spools, as shown in Figure 2, we derive the steady-state cylinder pressures in both cases.

3.4.1. Typical Pressure-Compensated Non-Differential Spool Valve

Firstly, for the asymmetric spool valve shown in Figure 2a, which is another typical pressure-compensated, hydro-mechanical valve, the turbulent flow equations are given by [29]:

$$Q_{\mathrm{A}} = \begin{cases} K_{\mathrm{vPA}}(I)\sqrt{\Delta p_{\mathrm{c}}}, & \text{if I} \geq 0 \text{ (cylinder extends)} \\ -K_{\mathrm{vAT}}(I)\sqrt{p_{\mathrm{A}} - p_{\mathrm{T}}}, & \text{if I} < 0 \text{ (cylinder retracts)} \end{cases} \tag{21}$$

$$Q_{\mathrm{B}} = \begin{cases} -K_{\mathrm{vBT}}(I)\sqrt{p_{\mathrm{B}} - p_{\mathrm{T}}}, & \text{if I} \geq 0 \text{ (cyl. ext.)} \\ K_{\mathrm{vPB}}(I)\sqrt{\Delta p_{\mathrm{c}}}, & \text{if I} < 0 \text{ (cyl. ret.)} \end{cases} \tag{22}$$

where I is the valve's control current, Δp_c is the constant pressure drop across the main spool due to the pressure compensation, p_{T} is the tank pressure and K_{vPA}, K_{vAT}, K_{vBT} and K_{vPB} are the flow coefficients of the P-A, A-T, B-T and P-B control edges, respectively. The flow coefficients $K_{\mathrm{vPA}}(I)$ and $K_{\mathrm{vBT}}(I)$ are monotonically increasing positive functions for a valve current I greater than a positive dead zone current $I_{\mathrm{DZ}}^{(+)}$. Otherwise, they have a value of zero. The flow coefficients $K_{\mathrm{vAT}}(I)$ and $K_{\mathrm{vPB}}(I)$ are similarly defined as monotonically increasing positive functions for a current I less than a negative dead zone current $I_{\mathrm{DZ}}^{(-)}$. Here, ideal pressure compensation was assumed, and reversed flow was neglected to simplify analysis.

Utilising Equations (20)–(22) and the external force (which equals the cylinder force $F = p_{\mathrm{ssA}}A_{\mathrm{A}} - p_{\mathrm{ssB}}A_{\mathrm{B}}$) to resolve the steady-state pressures yields [30]:

$$p_{\mathrm{ssA}} = \begin{cases} F/A_{\mathrm{A}} + \frac{A_{\mathrm{B}}}{A_{\mathrm{A}}}p_{\mathrm{ssB}}, & \text{for I} > I_{\mathrm{DZ}}^{(+)} \\ \underbrace{\left(\frac{K_{\mathrm{vPB}}A_{\mathrm{A}}}{K_{\mathrm{vAT}}A_{\mathrm{B}}}\right)^2}_{\text{Coeff.}}\Delta p_{\mathrm{c}} + p_{\mathrm{T}}, & \text{for I} < I_{\mathrm{DZ}}^{(-)} \end{cases} \tag{23}$$

$$p_{\mathrm{ssB}} = \begin{cases} \overbrace{\left(\frac{K_{\mathrm{vPA}}A_{\mathrm{B}}}{K_{\mathrm{vBT}}A_{\mathrm{A}}}\right)^2}^{\text{Coeff.}}\Delta p_{\mathrm{c}} + p_{\mathrm{T}}, & \text{for I} > I_{\mathrm{DZ}}^{(+)} \\ -F/A_{\mathrm{B}} + \frac{A_{\mathrm{A}}}{A_{\mathrm{B}}}p_{\mathrm{ssA}}, & \text{for I} < I_{\mathrm{DZ}}^{(-)} \end{cases} \tag{24}$$

Interestingly, the cylinder back-pressures can be approximated by multiplying the pressure compensator's constant pressure drop with a momentarily constant coefficient (labelled). These equations apply regardless of the load force direction. If the cylinder back-pressure is non-existent, as in single-acting lift cylinders, then these steady-state pressure approximations can be easily simplified.

3.4.2. Typical Pressure-Compensated Differential Spool Valve

The second spool type is the pressure-compensated differential spool valve (Figure 2b). Because of the spool's differential position, the fluid flows are regenerative from the cylinder chamber B to cylinder chamber A via the valve's P port in the cylinder extension case. This differential position reduces Q_{P}, the required pump flow rate, to $Q_{\mathrm{A}} - |Q_{\mathrm{B}}|$. Hence, the actuator flow rates can be written as:

$$Q_{\mathrm{A}} = \underbrace{K_{\mathrm{vPA}}(I)\sqrt{\Delta p_{\mathrm{c}}}}_{Q_{\mathrm{P}}} + \underbrace{|K_{\mathrm{vBP}}(I)|\sqrt{p_{\mathrm{B}} - p_{\mathrm{A}} - \Delta p_{\mathrm{c}}}}_{|Q_{\mathrm{B}}|} \tag{25}$$

$$Q_{\mathrm{B}} = -K_{\mathrm{vBP}}(I)\sqrt{p_{\mathrm{B}} - p_{\mathrm{A}} - \Delta p_{\mathrm{c}}} \tag{26}$$

where $\mathrm{I} > 0$ and K_{vPA} and K_{vBP} are defined as positive functions of current, Δp_{c} is a constant-pressure drop over the control edge P-A due to the pressure compensation and the sum of p_{A} and Δp_{c} equals the pressure in the valve's P port, omitting some pressure losses (Q_{B} denotes the flow rate over the control edge B-P). Using Equations (20), (25) and (26), the steady-state cylinder chamber pressures can be resolved, yielding:

$$p_{\mathrm{ssB}} = \frac{F}{A_{\mathrm{A}} - A_{\mathrm{B}}} + \left(\frac{K_{\mathrm{vPA}}^2 A_{\mathrm{B}}^2 A_{\mathrm{A}}}{K_{\mathrm{vPB}}^2 (A_{\mathrm{A}} - A_{\mathrm{B}})^3} + \frac{A_{\mathrm{A}}}{A_{\mathrm{A}} - A_{\mathrm{B}}} \right) \Delta p_{\mathrm{c}} \tag{27}$$

$$p_{\mathrm{ssA}} = F / A_{\mathrm{A}} + p_{\mathrm{ssB}} \frac{A_{\mathrm{B}}}{A_{\mathrm{A}}} \tag{28}$$

Here, the flow coefficients and the external force are again momentarily constant. Pressure losses in pipes, hoses and filters were omitted. In contrast to the non-differential case, the cylinder back-pressure p_{ssB} depends on the cylinder force. None of these derived steady-state pressures, however, depend explicitly on the supply pressure because of the pressure compensation.

4. Problem Formulation

Let us formulate our inherently continuous-time optimal control problem in discrete-time for numerical solution. By dividing the fixed optimisation time t_{f}, which corresponds to the desired duration of the end-effector path, into N intervals of length t_{f}/N, our optimal control problem can be expressed in discrete-time as follows:

$$\min_{\pi_u \in \pi_{\mathbb{U}}} \sum_{k=0}^{N-1} P_{\mathrm{h},k}(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k) + \ell_k(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k) \tag{29}$$

subject to:

$$\bar{\mathbf{x}}_{k+1} = \bar{\mathbf{F}}_k = \bar{\mathbf{x}}_k + T_{\mathrm{s}}, \mathbf{F}_k(\bar{\mathbf{x}}_k, \bar{u}_k) \tag{30}$$

where π_u denotes the control policy and $\pi_{\mathbb{U}}$ denotes the admissible control policies. The hydraulic power $P_{\mathrm{h},k}$, defined at each time stage $k \in \{0, 1, 2, \ldots, N-1\}$, is the principal cost to be minimised, and the term ℓ_k, which can be replaced by standard inequality constraints, penalises violations of actuator limits. For the performance cost $P_{\mathrm{h},k}$, effective functions which minimise hydraulic energy consumption are proposed because the conventional cost functions may be sub-optimal.

In the above, the system dynamics have been discretised using the Euler method and T_{s} as the integration step corresponding to the interval length t_{f}/N. The state vector $\bar{\mathbf{x}}_k = [\bar{c}_{4,k}\ \bar{v}_{4,k}]^{\mathrm{T}}$, with initial state $\bar{\mathbf{x}}_0$, contains the extension cylinder position and velocity. The extension cylinder acceleration $\bar{a}_{4,k}$ is used as the control $\bar{u}_k$. To optimise movements of the extension cylinder at the acceleration level, we create a double integrator system $\mathbf{F}_k(\bar{\mathbf{x}}_k, \bar{u}_k)$, denoted with:

$$\mathbf{F}_k(\bar{\mathbf{x}}_k, \bar{u}_k) = \begin{bmatrix} 0 & v_{4,\max}/c_{4,\max} \\ 0 & 0 \end{bmatrix} \bar{\mathbf{x}}_k + \begin{bmatrix} 0 \\ a_{4,\max}/v_{4,\max} \end{bmatrix} \bar{u}_k \tag{31}$$

where the maximum extension cylinder position, velocity and acceleration, $c_{4,\max}$, $v_{4,\max}$ and $a_{4,\max}$, respectively, have been applied to normalise the physical units for the cylinder's position, velocity and acceleration. These normalised units can be transformed into conventional physical units by employing scaling equations $\mathbf{x}_k = \mathrm{diag}(c_{4,\max}, v_{4,\max})\bar{\mathbf{x}}_k$ and $u_k = a_{4,\max}\bar{u}_k$. These extension cylinder coordinates can also be converted into joint coordinates by multiplying the cylinder motion by $1/r_4$, in view of Equations (10)–(12). This final transformation is required for the next stage.

4.1. Using Inverse Kinematics to Resolve the Motion of Non-Redundant Joints

Because our convenient method directly provides only the extension cylinder and joint motions for the cost functions, the method we propose includes the application of inverse kinematics to resolve the motions of the non-redundant joints. These motions are included in the vector:

$$\mathbf{x}_k^{(\mathrm{r})} = \begin{bmatrix} \mathbf{q}_k^{(\mathrm{r})} & \dot{\mathbf{q}}_k^{(\mathrm{r})} & \ddot{\mathbf{q}}_k^{(\mathrm{r})} \end{bmatrix}^{\mathrm{T}} \tag{32}$$

For example, the positions of the base, lift and tilt joints (the non-redundant joints) are denoted with $\mathbf{q}_k^{(\mathrm{r})} = [q_{1,k}\ q_{2,k}\ q_{3,k}]^{\mathrm{T}}$ at time stage k. Because the extension joint's motion is known, these non-redundant joint positions can be resolved directly by inverting Equations (4)–(6), which yields [27]:

$$\mathbf{q}_k^{(\mathrm{r})} = \begin{bmatrix} \arctan(z_{\mathrm{wd,k}}/x_{\mathrm{wd,k}}) & h_2 & h_3 \end{bmatrix}^{\mathrm{T}} \tag{33}$$

Here, $z_{\mathrm{wd},k}$ and $x_{\mathrm{wd},k}$ denote the components of the end-effector's desired position $\mathbf{x}_{\mathrm{td}}$ at time stage k in the direction of the z- and x-axes, respectively, in the global frame. The lengthy inverse kinematic functions h_2 and h_3 are omitted for brevity.

At the velocity level, the approach requires that we subtract the contribution of the extension joint from the desired end-effector velocity $\dot{\mathbf{x}}_{\mathrm{td},k}$. Then, the joint velocities in $\dot{\mathbf{q}}_k^{(\mathrm{r})}$ can be obtained by inverting Equation (7) as in:

$$\dot{\mathbf{q}}_k^{(\mathrm{r})} = \begin{bmatrix} \mathbf{J}_1(\mathbf{q}_k) & \mathbf{J}_2(\mathbf{q}_k) & \mathbf{J}_3(\mathbf{q}_k) \end{bmatrix}^{-1} \left(\dot{\mathbf{x}}_{\mathrm{td},k} - \mathbf{J}_4(\mathbf{q}_k)\dot{q}_{4,k} \right) \tag{34}$$

because the partial Jacobian inverse matrix $\left[\mathbf{J}_1(\mathbf{q}_k)\ \mathbf{J}_2(\mathbf{q}_k)\ \mathbf{J}_3(\mathbf{q}_k)\right]^{-1}$ is non-singular. This matrix can be symbolically computed in advance to possibly decrease the computational complexity.

Similarly, accelerations of the non-redundant joint can be resolved by inverting Equation (8):

$$\ddot{\mathbf{q}}_k^{(\mathrm{r})} = \left[\mathbf{J}_1(\mathbf{q}_k)\ \mathbf{J}_2(\mathbf{q}_k)\ \mathbf{J}_3(\mathbf{q}_k) \right]^{-1} \left(\ddot{\mathbf{x}}_{\mathrm{td},k} - \dot{\mathbf{J}}(\mathbf{q}_k, \dot{\mathbf{q}}_k)\dot{\mathbf{q}}_k - \mathbf{J}_4(\mathbf{q}_k)\ddot{q}_4 \right) \tag{35}$$

where the right-hand term in the parentheses contains the Cartesian acceleration required from the non-redundant joints to maintain the desired end-effector acceleration $\ddot{\mathbf{x}}_{\mathrm{td},k}$. The Jacobian time derivative $\dot{\mathbf{J}}(\mathbf{q}, \dot{\mathbf{q}})$ can be symbolically computed in advance using Equation (9). We are now in a position to apply the cylinder coordinate transformations presented to obtain the necessary motions of the cylinders of the non-redundant joints for the proposed cost functions.

The proposed method optimises the dynamical motions of all the cylinders. However, our method reduces computational complexity by only computing the necessary extension cylinder motion to optimise the dynamical motions of all the cylinders. This is possible because the motion of the extension cylinder in the 4-DOF manipulator determines the motion of the base, lift and tilt cylinders along a certain end-effector path. We could have also chosen to directly optimise the motion of the lift or tilt cylinder and use the motion of that cylinder to compute the required optimal dynamical motions of the base, lift/tilt and extension joints for the cost functions.

4.2. Proposed Cost Function: Power Consumption of Load-Sensing System

We first propose the power consumption the LS system as the principal cost in view of Equation (29) to be minimised. The cost function of the LS system, at time stage k, can be written as:

$$P_{\mathrm{h},k} = \frac{p_{\mathrm{s},k} Q_{\mathrm{p},k}}{\eta_k} \tag{36}$$

where $p_{\mathrm{s},k}$ denotes the pump supply pressure, $Q_{\mathrm{p},k}$ is the pump flow rate and η_k is the total efficiency of the hydraulic pump and the driving motor, which depends on the operating point, i.e.,

the displacement, supply pressure and rotational speed of the pump. Thus, the energy consumption of the LS system can be optimised by minimising the hydraulic output power or by increasing the driving system's efficiency. Since pump and motor efficiency data is not usually available, without loss of generality, we assume the total efficiency is constant.

The pump flow rate, which is a positive function without energy recuperation, can be approximated at time stage k by summing the cylinder flow rates as in:

$$Q_{\mathrm{p},k} = \sum_{i=1}^{4} v_{i,k}\left[A_{\mathrm{A}i}H(v_{i,k}) - A_{\mathrm{B}i}H(-v_{i,k})\right] \tag{37}$$

where $v_{i,k}$ is the velocity of cylinder i and $H(v_{i,k})$ denotes the piecewise Heaviside step function. The piston-side area is denoted with $A_{\mathrm{A}i}$ and the rod-side area with $A_{\mathrm{B}i}$. For simplicity, the hydraulic fluid is assumed to be incompressible. The rod-side area is negatively signed to enforce a positive pump flow rate when the cylinder is retracting and has negative velocity. The Heaviside step function can be defined using:

$$H(v_{i,k}) = \begin{cases} 0 & \text{if } v_{i,k} < 0 \\ \frac{1}{2} & \text{if } v_{i,k} = 0 \\ 1 & \text{if } v_{i,k} > 0 \end{cases} \tag{38}$$

which introduces a discontinuity in the objective function. Because having a differentiable objective function is desirable, the discontinuous Heaviside step function can be approximated by a differentiable expression, after some mathematical manipulations:

$$H(v_{i,k}) \approx \hat{H}(v_{i,k}) = \frac{1}{1 + e^{-2sv_{i,k}}} \tag{39}$$

where s is a real positive number, such as 20,000.

The pump supply pressure, according to the LS system's operating principle, can be calculated at steady-state at time stage k by using the maximum driven actuator pressure:

$$p_{\mathrm{s},k} = \max\left\{p_{1,k}, p_{2,k}, p_{3,k}, p_{4,k}\right\} + \Delta p_{\mathrm{LS}} \tag{40}$$

where $p_{1,k}$, $p_{2,k}$, $p_{3,k}$ and $p_{4,k}$ denote the chamber pressures of the base, lift, tilt and extension cylinders, respectively. The LS pressure margin Δp_{LS} is roughly 3 MPa in the test-bed.

The chamber pressure $p_{i,k}$ of cylinder i is taken from the rod- or piston-side depending on the cylinder's direction of motion. The pressure is zero if the actuator is not driven. Hence, the chamber pressure $p_{i,k}$ at time stage k can be estimated by using:

$$p_i = \left(p_{\mathrm{ssA},i}H(v_i - \varepsilon) + p_{\mathrm{ssB},i}H(-v_i - \varepsilon)\right)H(F_i v_i) \tag{41}$$

where the use of the minor positive constant ε in the Heaviside step functions assures zero chamber pressure when the cylinder is not driven and its velocity $v_i \in [-\varepsilon, \varepsilon]$. Time indices were omitted for clarity. In addition, multiplying by the Heaviside value of the actuator power ensures a non-negative actuator pressure when the cylinder load is over-running. In this over-running load case, the cylinder velocity v_i and the cylinder pressure force F_i have opposite signs. The steady-state pressure equations are chosen according to the spool type used either from Equations (23)–(24) or Equations (27)–(28).

A discontinuous maximum function has been used in Equation (40). This function can be approximated by applying the equation:

$$\max(p_1, p_2) = 0.5\left(p_1 + p_2 + |p_1 - p_2|\right) \tag{42}$$

thrice; i.e., $\max\{p_{1,k}, p_{2,k}, p_{3,k}, p_{4,k}\}$ can be computed as $\max(\max(\max(p_{1,k}, p_{2,k}), p_{3,k}), p_{4,k})$. The absolute function, in turn, can be approximated as the square root of $(p_1 - p_2)^2 + \epsilon$. This ϵ should

be a minor positive constant, such as 10^{-6}, to improve the accuracy of the approximation. As a result of these approximations, the LS cost is desirably continuously differentiable, which allows us to more easily employ a standard optimal control solver for performance comparison. The pump flow rate and supply pressure modelling results achieved with the test-bed can be found from Appendix C.

4.3. Proposed Cost Function: Power Consumption of Constant-Pressure System

In our second application, we minimise the energy consumption of the CP system. The cost function of the CP system can be written as:

$$P_{\mathrm{h},k} = \frac{p_{\mathrm{s}} Q_{\mathrm{p},k}}{\eta_k} \tag{43}$$

where p_{s} denotes the constant supply pressure and the pump flow rate $Q_{\mathrm{p},k}$ is given by Equation (37). Because of the constant pressure, the energy consumption of the CP system can generally be optimised load-independently without knowledge of manipulator dynamics through minimisation of the pump flow rate. By minimising the pump flow rate, the total volumes displaced by the hydraulic cylinders in the solution end up being minimised. Because the piston and piston-rod-side areas obviously affect the cylinder volumes displaced, minimising the energy consumption of the CP system does not equal to the actuator velocity norm minimisation, e.g., in [10]. The variation of cylinder areas among the manipulator's hydraulic cylinders, including the different piston and rod-side areas of these conventionally asymmetric cylinders, render the pump flow rate minimisation a highly tractable objective.

4.4. Actuator Limits

Because of the inherent physical limitations of the manipulator's cylinders, the cylinders' motions are limited. For this reason, we have introduced a mechanism through the optimal control problem's term $\ell_k(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k)$ to impose actuator constraints at the position, velocity and acceleration levels by issuing a significant penalty compared to the magnitude of the principal cost if the limits are violated.

Since the position limits depend on the cylinder stroke, they can be expressed effectively either in the joint or actuator space. In contrast, however, the velocity limits should be expressed in the actuator space because these limits cannot be expressed optimally from the perspective of the hydraulic system in the joint space. Based on the influential dimensions of the hydraulic pump and cylinders and the cylinders' non-constant torque arms these velocity limits are constant only in the actuator space to ensure that the total fluid flow rates for the actuator do not exceed the pump's maximum flow output. Assuming fluid incompressibility in these computations is the most convenient approach. Notice that by assuming constant velocity limits, the joint velocity limits vary if the torque arm is non-constant (see Equation (11)). Only if a particular cylinder's torque arm is constant can the corresponding velocity limit be expressed optimally in the joint space. Hence, in the test-bed case, only the velocity limits for base and extension cylinders can be expressed in either space.

Lastly, we choose constant acceleration limits for the cylinders for simplicity. Consequently, the joint acceleration limits vary (see Equation (12)). Selecting the proper acceleration limit magnitudes are important from at least two perspectives: firstly, to limit the occurrence of higher accelerations that might lower the manipulator wear life and secondly, to assure higher controller tracking performance. The actuator limits used in the experiments can be found in Table 3.

Table 3. Actuator/joint limits of the test-bed manipulator.

i	Joint Pos.	Cyl. Vel. [m/s]	Cyl. Acc. [m^2/s]
1	$[-\pi, \pi]$ rad	$[-0.10, 0.10]$	$[-0.4, 0.4]$
2	$[-0.425, 1.097]$ rad	$[-0.15, 0.15]$	$[-0.4, 0.4]$
3	$[-2.75, -0.25]$ rad	$[-0.15, 0.15]$	$[-0.4, 0.4]$
4	$[0, 2 \times 1.85]$ m	$[-0.40, 0.40]$	$[-0.4, 0.4]$

5. Dynamic Programming Solution

The complex optimal control problem of the hydraulic manipulator is solved to global optimality using the discrete-time DP algorithm (see Algorithm 1 [31–33]). The main advantage of the algorithm is that it can produce this solution considerably more effectively than a brute-force method, whose computational cost grows exponentially with the time stages N. However, the main disadvantage of the algorithm is that its computational cost grows exponentially with the states and controls due to the curse of dimensionality. Hence, our low-dimensional optimal control problem that we have reduced to the bare essentials makes this global DP algorithm applicable. Recall that we have reduced our optimal control problem size to only two states and one control.

The effectiveness of the DP algorithm over the brute-force method results from the principle of optimality, which states that "an optimal policy has the property that whatever the initial state and initial decision (control) are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" [31]. The Bellman equation, which characterises this principle of optimality, can be written at time stage k, in view of our objective Equation (29), as:

$$J_k^\star(\bar{\mathbf{x}}_k) = \min_{\bar{u}_k \in \mathbb{U}} \left\{ P_{\mathrm{h},k}(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k) + \ell_k(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k) + J_{k+1}^\star(\bar{\mathbf{F}}_k(\bar{\mathbf{x}}_k, \bar{u}_k)) \right\} \tag{44}$$

where the sum $P_{\mathrm{h},k} + \ell_k$ represents the running cost and the vector $\bar{\mathbf{F}}_k$ represents the system dynamics (see Equations (30)–(31)). The minimum cost-to-go $J_k^\star$ is the minimum sum of the running costs from stage $k \in \{0, 1, \ldots, N-1\}$ to the final stage at $k = N-1$. This optimal cost-to-go $J_k^\star(\bar{\mathbf{x}}_k)$ is defined for each state vector combination such that $\bar{\mathbf{x}}_k \in \mathbf{X} = \{\mathbb{X}_1 \times \mathbb{X}_2\}$ and minimised over all possible controls at stage k such that $\bar{u}_k \in \mathbb{U}$. Here, $\mathbb{X}_i = \{x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(N_{\mathrm{x}_i})}\}$ and $\mathbb{U} = \{u^{(1)}, u^{(2)}, \ldots, u^{(N_{\mathrm{u}})}\}$ denote the discrete-state and control sets, respectively. The variables N_{x_i} and N_{u}, in turn, determine the density of the state and control grids, respectively. Sufficiently dense grids are defined to reach the global optimum. These grid sizes, and suitable time-discretisation, can be found through trial and error, taking note that the computational and memory requirements of the algorithm scale linearly with the density of the grids.

The algorithm's implementation shown in Algorithm 1 entails two consecutive phases: a backwards phase followed by a forward simulation. During the computationally heavier backwards phase, the optimal controls are resolved at each position in the state grid while the algorithm iterates backwards in time. In greater detail, at any given time stage k, the system is simulated (line 5), subsequent computations are performed to evaluate the cost function (lines 6–8), the minimum cost-to-go $J_k^\star(\bar{\mathbf{x}}_{k+1})$ computed at the previous iteration is interpolated because the system dynamics have likely transformed into a state unspecified in the grid (line 9) and the cost-to-go $J_k(\bar{\mathbf{x}}_k, \bar{\mathbf{u}}_k)$ is evaluated for all the states and controls by utilising the minimum cost-to-go $J_{k+1}^\star(\bar{\mathbf{x}}_{k+1})$ (line 10). From the cost-to-go evaluated at each control, the minimum cost-to-go $J_k^\star(\bar{\mathbf{x}}_k)$ is then stored for the next iteration (line 12) together with the optimal control $u_k^\star(\bar{\mathbf{x}}_k)$, which minimised the cost-to-go (line 13). The backwards phase iterates in this manner until it reaches the initial time stage. At this point, with our system of two states and one control, the algorithm has produced a two-dimensional state grid at each time stage, and the optimal controls have been stored at each position in the state grid. The algorithm then simulates in the forward direction from the initial time stage and state by following the grid of the optimal controls produced in the backwards phase. Interpolation of the optimal control is used during the forward simulation because is it is unlikely that the optimal controls were evaluated at the optimal states. When the simulation has been completed, the algorithm has produced the optimal control sequence and state trajectories as outputs. In addition to producing the global solution, this algorithm has the significant advantage that different initial configurations of the manipulator can be considered with just a forward simulation. This advantage is in contrast to most other optimisation algorithms where new optimisations from scratch are required.

Algorithm 1 Discrete-time dynamic programming algorithm

1. Backward phase with interpolation

Require: Sets $\mathbf{X}$ and $\mathbb{U}$
Initialisation: set final cost-to-go $J_N^\star(\bar{\mathbf{x}}_N)$ for all $\bar{\mathbf{x}}_N \in \mathbf{X}$
for $k = N-1$ to 0 **do**
  for all $\bar{\mathbf{x}}_k \in \mathbf{X}$ **do**
    for all $\bar{u}_k \in \mathbb{U}$ **do**
      $\bar{\mathbf{x}}_{k+1} = \bar{\mathbf{F}}_k(\bar{\mathbf{x}}_k, \bar{u}_k)$, see Equations (30) and (31).
      Determine $q_{4,k}$, $\dot{q}_{4,k}$ and $\ddot{q}_{4,k}$ using Equations (10)–(12).
      Complete $\mathbf{q}_k$, $\dot{\mathbf{q}}_k$ and $\ddot{\mathbf{q}}_k$ using Equations (33)–(35).
      Compute base, lift and tilt cyl. coordinates using Equations (10)–(12).
      Interpolate $J_{k+1}^\star(\bar{\mathbf{x}}_{k+1})$
      $J_k(\bar{\mathbf{x}}_k, \bar{\mathbf{u}}_k) = P_{\mathrm{h},k}(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k) + \ell_k(\mathbf{q}_k, \dot{\mathbf{q}}_k, \ddot{\mathbf{q}}_k) + J_{k+1}^\star(\bar{\mathbf{x}}_{k+1})$
    end for
    $J_k^\star(\bar{\mathbf{x}}_k) = \min_{\bar{u}_k \in \mathbb{U}} J_k(\bar{\mathbf{x}}_k, \bar{u}_k)$
    $u_k^\star(\bar{\mathbf{x}}_k) = \operatorname{argmin}_{\bar{u}_k \in \mathbb{U}} J_k(\bar{\mathbf{x}}_k, \bar{u}_k)$
  end for
end for

2. Forward simulation with interpolation

Require: $u_k^\star(\bar{\mathbf{x}}_k)$ for all $k \in \{0, 1, \ldots, N-1\}$ and $\bar{\mathbf{x}}_k \in \mathbf{X}$
Initialisation: set $\bar{\mathbf{x}}_0$
for $k = 0$ to $N-1$ **do**
  Interpolate $u_k^\star(\bar{\mathbf{x}}_k)$
  $\bar{\mathbf{x}}_{k+1} = \bar{\mathbf{F}}_k(\bar{\mathbf{x}}_k, u_k^\star(\bar{\mathbf{x}}_k))$
end for
Determine $q_{4,k}$, $\dot{q}_{4,k}$ and $\ddot{q}_{4,k}$ using Equations (10)–(12).
Complete $\mathbf{q}_k$, $\dot{\mathbf{q}}_k$ and $\ddot{\mathbf{q}}_k$ using Equations (33)–(35).

A demonstratively simple example of the DP algorithm has been presented, e.g., by Nurmi and Mattila [21]. The utilised general Matlab-coded DP algorithm can be found from Sundstrom and Guzzella [32]. Different DP algorithm varieties have been compared by van Berkel et al. [34] and Böhme and Frank [33], and a convergence proof of discrete-time DP has been presented by Shin and McKay [35].

6. Experimental Results

Before diving head-first into the experimental results from the real hydraulic manipulator introduced in Section 2, we describe the testing methodology. Firstly, the optimal control solutions compared consist of:

(1) the proposed LS cost (36), which is minimised by using the DP algorithm and named DP (LS) according to the naming convention "method (minimised cost)";
(2) the proposed LS cost, which is minimised by using the DIDO algorithm [22] and named DIDO (LS);
(3) the proposed CP cost (43), which is minimised according to DP (CP);
(4) the proposed CP cost, which is minimised according to DIDO (CP);
(5) a positive actuator energy cost, which is solved according to DP (A. En.) and which minimises $P_{\mathrm{h},k} = \sum_{i=1}^{4} F_{i,k} v_{i,k}$ for positive work, $F_{i,k} v_{i,k} > 0$, and
(6) an extension cylinder use cost, which is solved according to DP (ext.) and which minimises $P_{\mathrm{h},k} = v_{4,k}^2$.

Secondly, the instantaneously optimal solutions included in the comparison are

(7) the joint-limited pseudo-inverse solution by Flacco et al. [20] named P-inv. (j.), which minimises $\dot{\mathbf{q}}^{\mathrm{T}} \dot{\mathbf{q}}$, and

(8) an actuator-limited pseudo-inverse solution named P-inv. (a.), which was derived by combining the solutions of Beiner and Mattila [10] and Flacco et al. [20] to minimise $\mathbf{v}^{\mathrm{T}}\mathbf{v}$.

The DP solution (A. En.) is formulated with positive actuator work only to obtain the best attainable performance in the test-bed, which did not have the possibility of recovering energy. The solution for minimising use of the extension cylinder is included in the comparison to evaluate whether the solution lowers energy consumption, presumably, by reducing the lift and tilt cylinder forces required.

The DP algorithm was parametrised by dividing the state and control grid into 356 × 101 × 151 ($N_{x_1} \times N_{x_2} \times N_u$) discrete values. With this grid size and by using a time discretisation of 0.05 s, we could nearly max out the 8 GB of RAM available. In the evaluation, the performance of the DP algorithm is compared to a "bootstrapped" DIDO solver by employing the proposed problem formulation. This DIDO algorithm is a commercial, pseudospectral and Matlab-based optimal control solver that has been extensively used in the solution for aerospace optimal control problems.

The energy consumption of the LS and CP systems was computed as the performance indices to be compared by measuring the pump supply pressure and the flow rate along three kinds of end-effector paths. The end-effector was controlled without payload along these paths. Firstly, we present the energy consumption in the horizontal and vertical point-to-point paths where the end-effector moves from point A to point B without a payload in a straight line (see Figure 9a). Jazar's quintic rest-to-rest polynomial algorithm [36] was used to generate the Cartesian motion profile between these points. The horizontal path was driven in 10 s, requiring a maximum end-effector velocity of 0.60 m/s along the x_w-axis. The vertical path was driven in 6 s, and it required a higher maximum end-effector velocity of 0.95 m/s along the y_w-axis. Secondly, we consider half of the log-loading cycle where the end-effector starts from the load space (or trailer) of the forwarder at point A and moves the empty log grapple back to the logs on the ground at point B (see Figure 9b). A cubic spline was used to generate the desired end-effector path via points in this case (see e.g., Biagiotti and Melchiorri [37]). This path was driven in 15 s, and the maximum end-effector velocity was 0.75 m/s in the $x_w y_w z_w$-space.

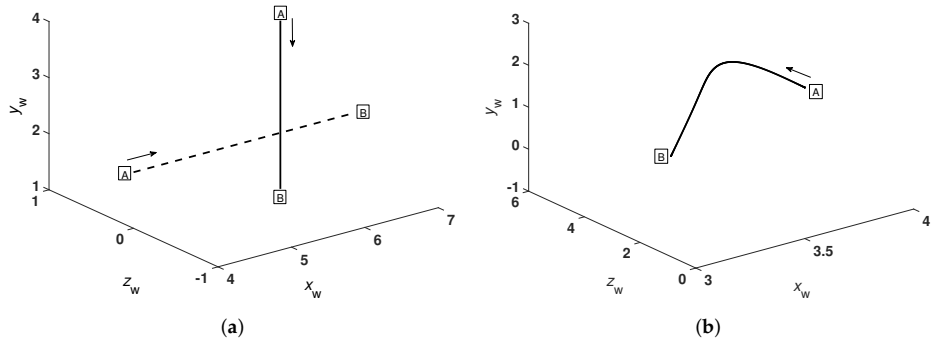


Figure 9. Test paths: (**a**) Horizontal and vertical paths; and (**b**) return half of log-loading cycle.

The paths were driven three times using each redundancy resolution from (1) through (8), and the results were averaged. Two initial configurations of the manipulator at the beginning of each path were also considered: firstly, based on the smallest feasible extension cylinder length and subsequently based on the middle point between the smallest and largest feasible extension cylinder lengths. The supply pressure of the CP system was set to 20 MPa in the experiments, and the motion controller used is discussed in Appendix A.

6.1. Horizontal Path

The experimentally obtained energy consumption of the LS system of each redundancy resolution has been sorted by magnitude in Figure 10a. These results demonstrate that the pseudo-inverses clearly have a sub-optimal performance by yielding 25% greater energy consumption than the best proposed hydraulic-energy optimal solution. The DIDO and DP solutions produce comparable results. In this path, minimising the energy consumption of the actuators is also similar to minimising the energy consumption of the LS system. Another important observation is the adequate performance of the CP solution, considering that it is based on a simple model. In addition, because the extension joint use minimising cost yields a poor solution, minimising the use of the extension joint appears to be impractical for our goal of minimising the hydraulic energy consumption.

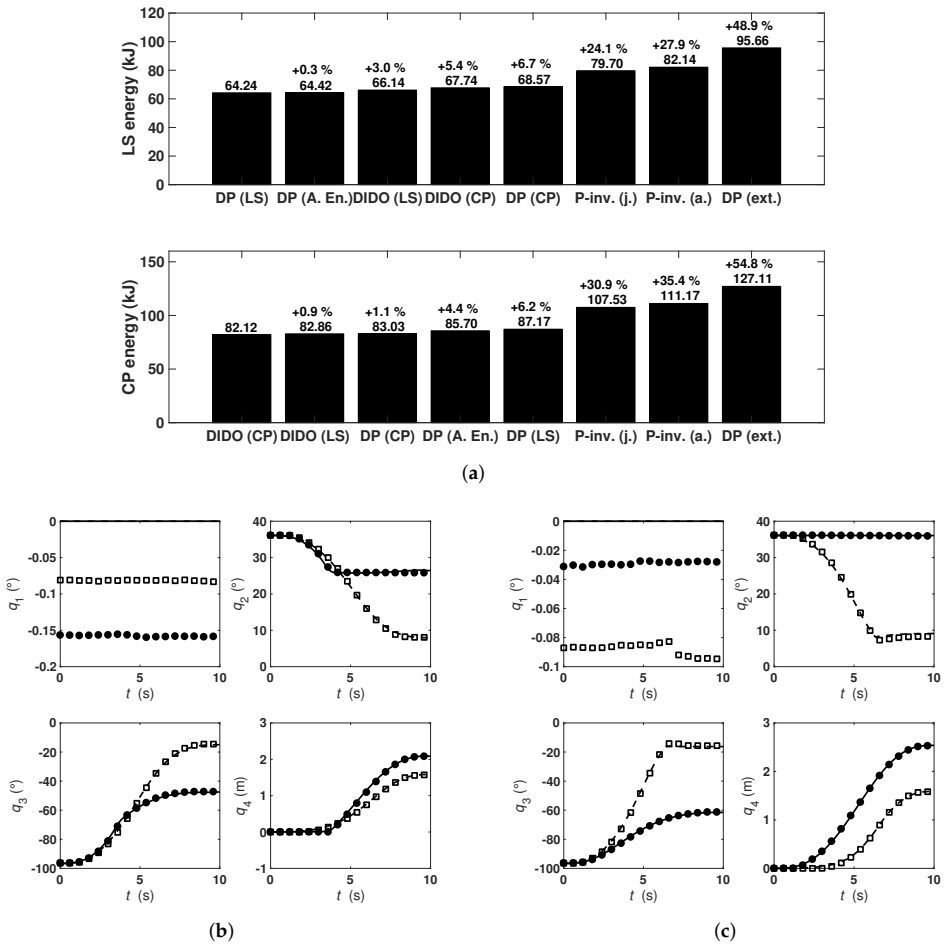


Figure 10. Horizontal path: (**a**) Experimental constant-pressure and load-sensing system energy consumptions (smallest feasible extension cylinder length at the initial configuration); (**b**) load-sensing system: energy-optimal DP (LS) joint trajectories (—) and controller tracking performance (●). Sub-optimal P-inv. (j.) joint trajectories (– –) and controller tracking performance (□); and (**c**) Constant-pressure system: energy-optimal DIDO (CP) joint trajectories (—) and controller tracking performance (●). Sub-optimal P-inv. (a.) joint trajectories (– –) and controller tracking performance (□).

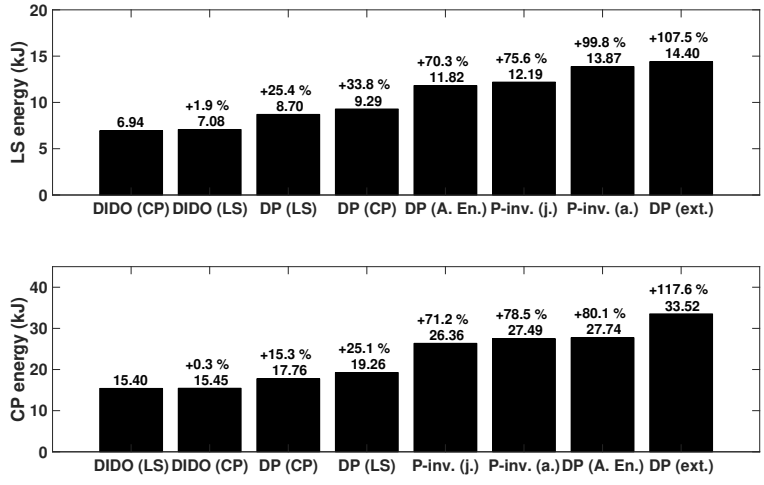
As can be seen in Figure 10a, the CP system results are comparable to the LS results, but the main difference is the greater energy consumption. The pseudo-inverses, again, have an inferior performance from the perspective of the energy consumption of the CP system, and they require approximately 30% greater energy use than do the hydraulic energy-optimal methods in this case. Inferiority of the pseudo-inverses reinforces the notion that minimising the velocities of the actuators does not minimise the energy consumption of the CP system. In addition, the best solutions here consume the same amount of CP energy as do the pseudo-inverses in the LS case with the same path.

Two of the redundancy resolutions' joint trajectories and controller performances have been illustrated in Figure 10b,c. The DP joint references, as shown there, were always low-pass filtered (cutoff frequency 5 Hz) to smooth out the trajectories' slightly jagged edges. This filtering caused no apparent error in the desired end-effector path. The results with the second initial configuration can be found in Appendix D (see Figure A4).

6.2. Vertical Path

Continuing with the results of the vertical path, we firstly compare the sorted LS energy-consumption presented in Figure 11a. As previously, the energy-optimised solutions for the hydraulic manipulators are superior compared to the pseudo-inverses, which require almost twice the energy consumption. However, the DP energy-optimised solutions for hydraulic manipulators have not reached a global optimum because the DIDO method clearly fares better in this case. The DIDO solutions themselves are practically the same. This path also demonstrates that minimising the energy consumption of the actuators does not always minimise the energy consumption of the LS system. The proposed solutions minimising the CP cost, again, have an adequate performance.

The energy-consumptions rates of the CP system presented in Figure 11a are similar, but the energy-consumption rates are approximately double compared to the LS case. The pseudo-inverses, most evidently, have an inferior performance compared to the proposed solutions, which, in contrast, require the lowest energy consumption.



(**a**)

Figure 11. *Cont.*

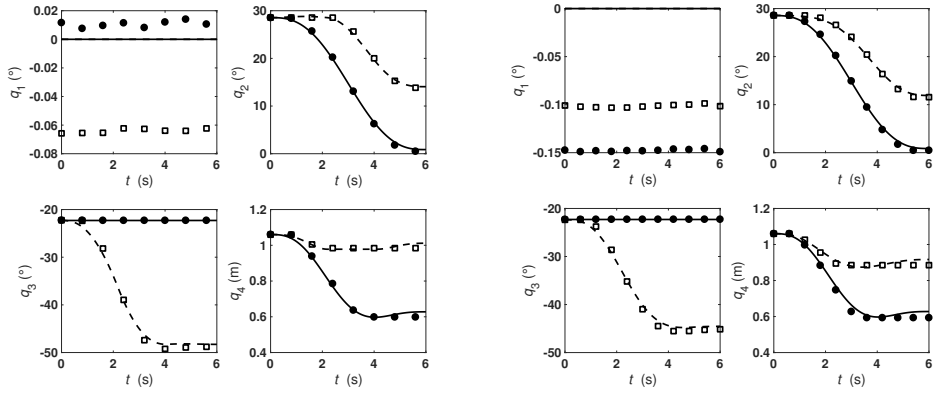


Figure 11. Vertical path: (**a**) Experimental constant-pressure and load-sensing system energy consumptions (smallest feasible extension cylinder length at the initial configuration); (**b**) load-sensing system: energy-optimal DIDO (CP) joint trajectories (—) and controller tracking performance (●). Sub-optimal DP (A. En.) joint trajectories (– –) and controller tracking performance (□); and (**c**) constant-pressure system: energy-optimal DIDO (LS) joint trajectories (—) and controller tracking performance (●). Sub-optimal P-inv. (a.) joint trajectories (– –) and controller tracking performance (□).

A few optimal and sub-optimal joint trajectories and controller performances are illustrated in Figure 11b,c. Compared to the previous path, the controller tracking performances are slightly worse because the end-effector moves faster. The energy consumption with the second initial configuration is presented in Figure A5 in Appendix D. These energy consumption rates demonstrate that the initial manipulator configuration has a significant impact on the energy consumption. Figure A5 also shows that the algorithms, or DP (CP) specifically in this case, can sometimes yield an average result. We suspect that this result is related to a combination of the DP algorithm's grid size, closed-loop dynamics and modelling errors.

6.3. Half-Cycle

Finally, we have the "half-cycle" energy consumption. The energy-consumption rates of the LS system presented in Figure 12a demonstrate that the redundancy resolutions' perform fairly similarly in this case. The minimisation of the actuator energy is also optimal in this case. The joint-limited pseudo-inverse, surprisingly, produces an energy-consumption rate comparable to the consumption obtained using hydraulic energy-optimised redundancy resolutions. Because the simulations (not presented in this paper) predicted a greater difference between these solutions, the smaller difference originates from modelling errors and the behaviour of the closed-loop control. The energy-consumption rates for the CP system presented in Figure 12a differ from the LS results in that the pseudo-inverses are inferior.

Selected optimal and sub-optimal joint trajectories and controller performances from the half-cycle are illustrated in Figure 12b,c. Furthermore, the controller performances of each test case were evaluated according to the ISO 9283 standard, which contains a standardised general method for computing the maximum end-effector path tracking error (see Table 4). The controller performances are superior when the CP hydraulic system is used, particularly over the half-cycle path. This is not surprising because the LS system's actuator flow dynamics are slower due to the non-constant pressure level. The tracking errors, in general, are sufficiently insignificant and consistent for evaluating the proposed solution. Again, the result with the second initial configuration can be found in Appendix D (Figure A6).

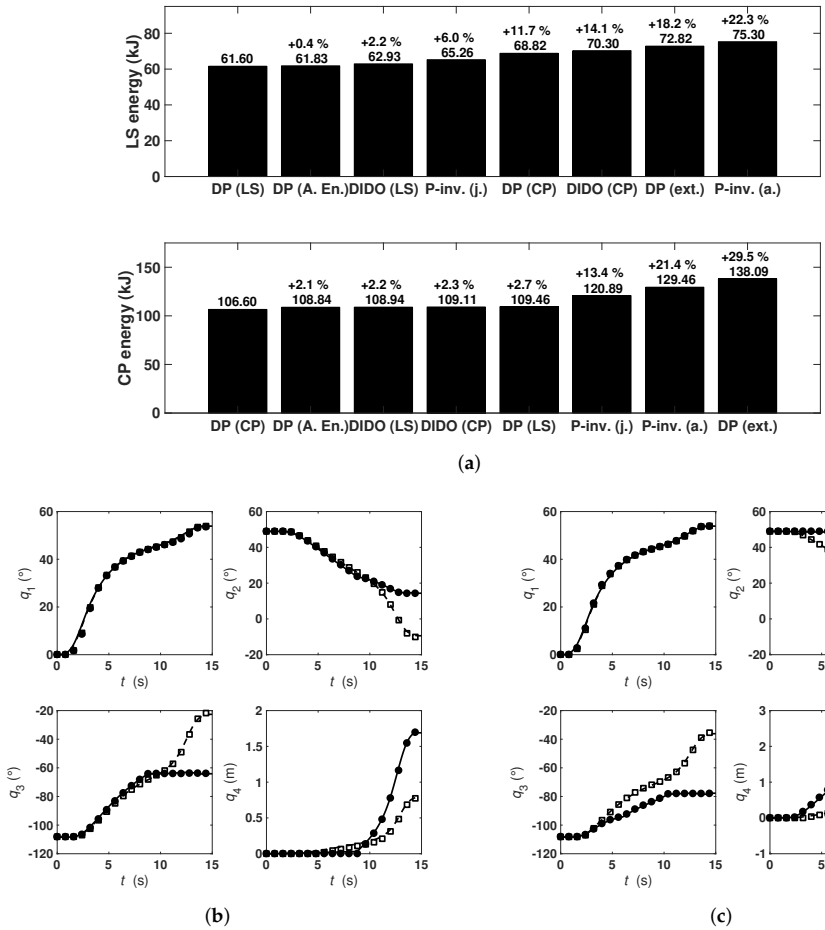


Figure 12. Half-cycle: (**a**) Experimental constant-pressure and load-sensing system energy consumptions (smallest feasible extension cylinder length at the initial configuration); (**b**) load-sensing system: energy-optimal DP (LS) joint trajectories (—) and controller tracking performance (●). Sub-optimal P-inv. (a.) joint trajectories (− −) and controller tracking performance (□); and (**c**) constant-pressure system: energy-optimal DP (CP) joint trajectories (—) and controller tracking performance (●). Sub-optimal P-inv. (j.) joint trajectories (− −) and controller tracking performance (□).

As a conclusion of the evaluation, the pseudo-inverses, in particular, and actuator energy solutions, in some cases, were inferior compared to the proposed hydraulic energy-optimised redundancy resolutions. These inferior solutions should be used with caution when minimising hydraulic energy consumption. The DIDO solutions were, for the most part, comparable to the DP solutions. The CP optimal solutions, in turn, were often comparable to the LS optimal solutions. The vertical and horizontal paths demonstrated the greatest energy savings, but energy savings were also attainable in the half-cycle when energy-optimised redundancy resolution was used. A solution containing both halves of the log-loading cycle could be obtained from a combination of a minimum-time solution for the log-loading part and the energy-optimal solution proposed for this return part. The results were obtained with the assumption of constant pumping efficiency. In the conditions of the experiments, in which the pump flow rate was typically less than 25% of the maximum flow rate of the pump, it might

have been advantageous to incorporate the variations in pumping efficiency in the optimisation to improve the results. However, the satisfactory results of our paper suggest that it might not be worthwhile to consider the efficiency map of the pump in the optimisation.

Table 4. Maximum end-effector path tracking errors in each redundancy resolution computed according to ISO 9283 standard [38].

Path	Solution	Track. Err. (m)	
		Load-Sensing System	**Constant-Pressure System**
Horizontal	DP (CP)	0.088	0.070
	DP (LS)	0.076	0.073
	DP (A. En.)	0.064	0.064
	DP (ext.)	0.146	0.140
	P-inv. (a.)	0.127	0.069
	P-inv. (j.)	0.057	0.052
	DIDO (CP)	0.064	0.041
	DIDO (LS)	0.063	0.049
Vertical	DP (CP)	0.098	0.091
	DP (LS)	0.093	0.087
	DP (A. En.)	0.076	0.085
	DP (ext.)	0.085	0.086
	P-inv. (a.)	0.108	0.095
	P-inv. (j.)	0.098	0.101
	DIDO (CP)	0.077	0.080
	DIDO (LS)	0.079	0.081
Half-cycle	DP (CP)	0.346	0.129
	DP (LS)	0.265	0.117
	DP (A. En.)	0.232	0.111
	DP (ext.)	0.210	0.155
	P-inv. (a.)	0.200	0.219
	P-inv. (j.)	0.343	0.207
	DIDO (CP)	0.277	0.148
	DIDO (LS)	0.245	0.179

7. Conclusions and Future Work

This paper addressed the open problem of global energy-optimal redundancy resolution of 4-DOF hydraulic manipulators in robotic control applications, in which the highly repetitive manipulator movements are automated to improve work productivity and operator workload circumstances. To ensure the hydraulic energy-optimality of the redundancy resolution, the solution was formulated at the hydraulic system level by proposing effective cost functions for the commonly used LS and CP hydraulic systems. Furthermore, actuator constraints at the position, velocity and acceleration levels were imposed in the solution. To reduce the computational complexity and problem dimensionality, firstly, only the motion of the redundant extension cylinder was directly optimised. Secondly, hydraulic system properties were incorporated through steady-state equations in the proposed cost functions. The non-redundant actuator motions were resolved via non-singular inverse kinematics to obtain the full manipulator motion state for the cost functions. The global DP algorithm was applicable to our optimal control problem formulation due to the conveniently reduced problem dimensionality. A pseudocode implementation of the DP algorithm used in the optimal control problem solution was presented to discuss the properties of the algorithm and to showcase its relatively simple implementation. Field experiments that were performed on a 4-DOF forestry manipulator demonstrated the significant advantage of the new global optimal control problem solution. Around 15–30% greater hydraulic energy consumption was observed with the conventional solutions in the LS and CP systems. These experiments also showed that the CP system's cost function for basic energy consumption, most often, also led to minimised energy consumption by the LS

system. Overall, these experimental results demonstrate energy-optimal redundancy resolution at the hydraulic system level in the prospective automated robotic applications. The existing practices in redundancy resolution of hydraulic manipulators are, therefore, significantly expanded. Experimental evaluation of the solution performance with an end-effector load was left for future work to limit the study's scope.

Acknowledgments: This work was funded by the Academy of Finland under project AUTOGRACE, Decision No. 304604. This funding is highly appreciated. The authors are also grateful for the technical help provided by Janne Honkakorpi, Janne Tuominen, Ari Moskari, Soheil Zavari, Anssi Mäenpää and Pauli Mustalahti.

Author Contributions: Jarmo Nurmi developed the redundancy resolution; Jarmo Nurmi conceived, designed and performed the experiments; Jarmo Nurmi also analysed the data to validate the redundancy resolution. Both authors contributed to the writing of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CAN	Controller Area Network
CP	Constant-pressure
DH	Denavit-Hartenberg
DOF	Degrees-of-freedom
DP	Dynamic programming
ISO	International Organization for Standardization
LS	Load-sensing

Appendix A. Motion Controller

A motion controller for each of the hydraulic cylinders was designed to evaluate the proposed redundancy resolution. This meant using the mobile hydraulic valve and the joint encoders of the test-bed in the motion control. This mobile hydraulic valve had a low bandwidth (less than 10 Hz), significant dead-zone (around 53%) and a time-delay greater than 0.15 s. In addition, the time-delay varied because of the hydraulic system dynamics. Therefore, only an adequate position tracking accuracy for the purposes of the evaluation, also considering the basic instrumentation used, was expected.

We designed the motion controller based on a standard feed-forward and dead-zone compensation strategy, where the valve control is inferred from an inverted table. This table compensates for the valve's dead-zone nonlinearity by matching a valve control signal to a given cylinder velocity (see Figure A1). By omitting the sub-indices denoting the actuator, we can write

$$I = DZI(v_\mathrm{d}) \tag{A1}$$

where v_d is the desired cylinder velocity and DZI denotes the compensating feed-forward function. This dead-zone compensation is crucial because an uncompensated dead-zone would significantly increase the control error [39]. Since the feed-forward is a static representation of the dynamical system, the controller input is augmented by a P-controller, i.e.,

$$I = DZI(v_\mathrm{d} + K_\mathrm{p} e) \tag{A2}$$

to add robustness to disturbances. Here, K_p is the P-gain and e is the cylinder position error. Employing this type of controller is standard practice, and it has been found to be surprisingly effective when used in conjunction with a mobile hydraulic valve [30,40]. This effectiveness arises from the dead-zone compensation and from the fact that the pressure-compensated mobile hydraulic valve inherently attenuates the effects of load disturbances on the actuator flows. Thus, a representative cylinder velocity and valve current table can be found regardless of the payload magnitude. Most often, however,

the mobile hydraulic valves are equipped with force-feedback spools. For this reason, a combination of the feed-forward and closed-loop control is practically necessary.

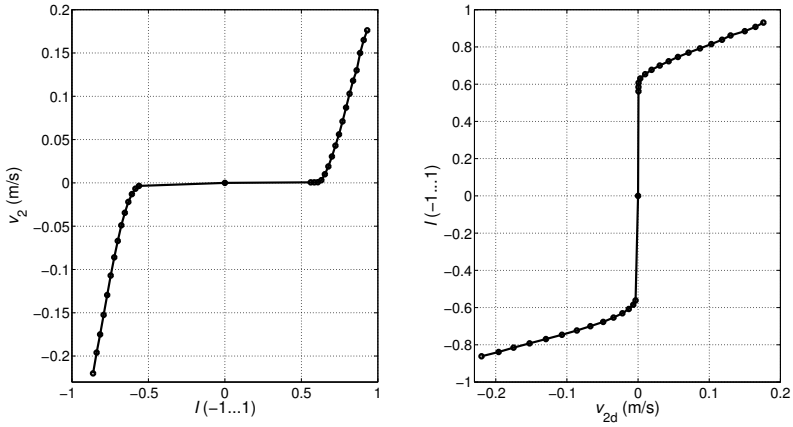


Figure A1. Determination of the feed-forward function: the measured lift velocities versus the valve control are illustrated on the left, the inverted feed-forward function DZI is illustrated on the right. The data point pairs can be matched by using a constant valve input to determine the steady-state cylinder velocity. The cylinder velocities in between the measured points can be interpolated using splines.

Another useful control practice we used was limiting controller activation to avoid unnecessary energy consumption and the signalling of the pump pressure reference through the LS pilot line the of mobile hydraulic valve. We implemented this strategy so that if the absolute value of the cylinder velocity tracking error was less than a minor positive constant ε_v, the mobile hydraulic valve would not be controlled (tracking mode). However, if the absolute value of the cylinder position error was greater than a minor positive constant ε_p, the controller would activate (regulation mode). Thus, if a minor cylinder velocity were desired, for example due to a numerical imprecision of the optimisation, the cylinder would not be forced to move. This controller is simpler than the nonlinear model-based controllers, for example, by Bu and Yao [41] and Koivumäki and Mattila [42].

Appendix B. Cylinder Forces

The supply pressure computations are based on accurate cylinder force models. The cylinder forces, as given by Equation (16), were identified by using a least squares algorithm. However, because the base cylinder was not subjected to a gravitational force on the level ground of our test area, we approximated the base cylinder force as a constant and focused on the other force models. The modelling errors of the cylinder forces are acceptable, as can be verified from Figure A2a–c.

To achieve the modelling accuracy, we paid attention to the modelling of the friction force of the extension function, which had a high-magnitude compared to the gravitational component. Moreover, the friction force exhibited steps because of the long and slightly flexible telescoping link; see Nielsen [43]. Thus, an unconventional friction force model

$$\begin{aligned} F_{f4} &= \tanh(K\dot{q}_4)F_{c4\text{total}} \\ &= \tanh(K\dot{q}_4)\big(F_{c4} + H_{f1}F_{c4}(q_4 - q_{f1}) + H_{f2}F_{c4}(q_{f2} - q_{f1}) \\ &\quad + H_{f3}F_{c4}(q_4 - q_{f3}) + H_{f4}F_{c4}(q_{f4} - q_{f3})\big) \end{aligned} \tag{A3}$$

was fitted to the data. Here, the tanh function, with a constant K of 2000, was used to assure smooth transient over the zero velocity, F_{c4} is the Coulomb friction force and $F_{c4total}$ is the total friction force (see Figure A2d). The Heaviside step functions

$$\begin{aligned} H_{f1} &= H(q_4 - q_{f1}) + H(q_{f2} - q_4) - 1, \\ H_{f2} &= H(q_4 - q_{f2}), \\ H_{f3} &= H(q_4 - q_{f3}) + H(q_{f4} - q_4) - 1 \text{ and} \\ H_{f4} &= H(q_4 - q_{f4}) \end{aligned} \tag{A4}$$

were chosen to obtain the behaviour with the steps and plateaus occurring at the positions q_{f1}, q_{f2}, q_{f3} and q_{f4}.

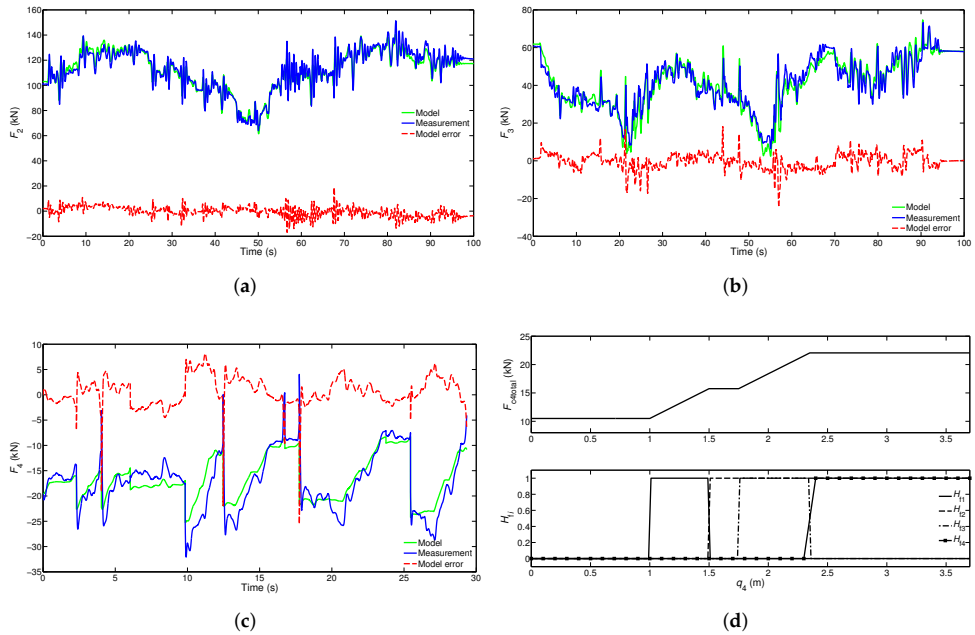


Figure A2. Cylinder forces F_i (kN) as a function of time (s): (**a**) Lift cylinder force F_2; (**b**) tilt cylinder force F_3; and (**c**) extension cylinder force F_4. The designed extension friction force $F_{c4total}$, where the Heaviside functions create the steps, is shown in (**d**).

Appendix C. Pump Flow Rate and Supply Pressure

The key pump flow rate and supply pressure modelling results and parametrisation of the test-bed are presented in the following. The pump flow rate model, as given by Equation (37), was parametrised with the cylinder dimensions, as given in Table 1. However,

(1) the rod-side area A_{B2} of the lift cylinder was set to zero since the cylinder's rod-side chamber was connected to the tank and
(2) the piston-side area A_{A4} of the extension cylinder was set to $A_{A4} - A_{B4}$ because the fluid flow rate circulated from the metre-out to the metre-in side.

Moreover, a simple empirical polynomial model, which adjusted "tilt cylinder area" as a function of cylinder velocity when the cylinder load was over-running, was used for the anti-cavitation of the tilt cylinder.

After improving modelling accuracy by taking these measures, the modelled pump fluid flow rate matches the measured flow rate dynamics well (see the uppermost sub-plot of Figure A3). The minor modelling errors are caused by the assumption of the incompressible fluid and the steady-state errors produced by the empirical anti-cavitation model of the tilt cylinder when this model affects the predictions, for example, at 21 and 78 s. However, the modelling errors at these time instants would be significant without the use of the anti-cavitation model.

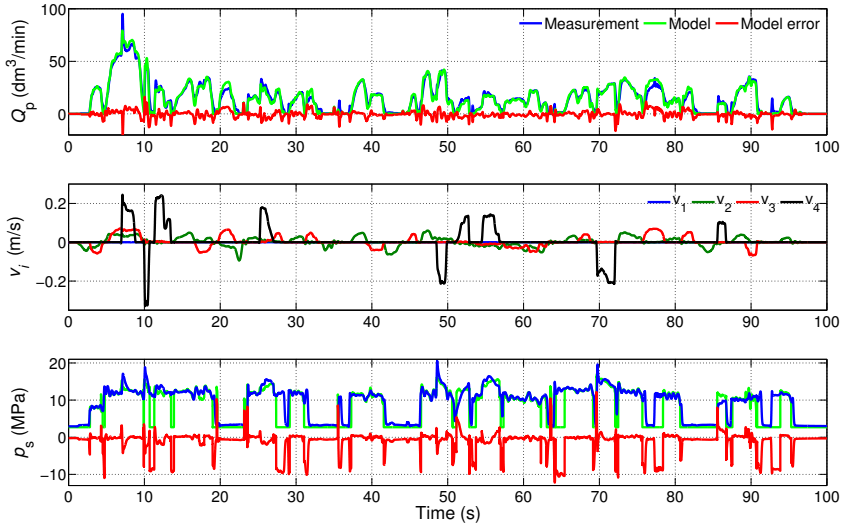


Figure A3. Pump flow rate Q_P (dm^3/min), cylinder velocities v_i (m/s) and pump supply pressure p_s (MPa) as a function of time (s).

The supply pressure model of the pump, as given by Equations (40)–(41), was parametrised using

(1) 0.005 m/s for the constant ε of Equation (41) in the model verification, but $\varepsilon = 0.002$ m/s for the optimisation;
(2) a zero tank pressure p_T;
(3) constant flow coefficients and
(4) pressure-compensated flow rate of the extension cylinder. Although the flow rate of the extension cylinder was not pressure-compensated by a dedicated valve like in the other cylinders, we could use the steady-state pressure model derived for the pressure-compensated differential case because the LS pump operated like a dedicated pressure compensator for the extension cylinder's high-pressure level.

As can be seen in the lowermost sub-plot of Figure A3, the modelled supply pressure tracks the measured pressure reasonably well (compared to the results presented in [44]). The observed modelling errors are caused mainly by the omitted pressure dynamics, but certain discrepancies originate from the residual oscillation of the input cylinder velocities around the zero velocity, which the model mistakes for a driven velocity, despite the value of $\varepsilon = 0.005$ m/s. Since modelling errors originating from this residual oscillation do not affect the optimisation, we used the smaller value for the ε in the optimisation.

Appendix D. Energy Consumption Using the Second Initial Configuration of the Manipulator

The following are experimental energy consumptions from the real test-bed using the second initial configuration of the manipulator at the beginning of each path. This configuration was computed based on the middle point between the smallest and largest feasible extension cylinder lengths.

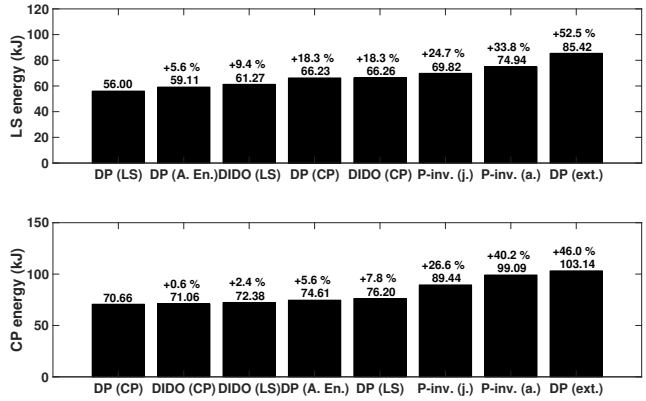


Figure A4. Horizontal path energy consumption replicated using middle feasible extension cylinder length at the beginning of the path.

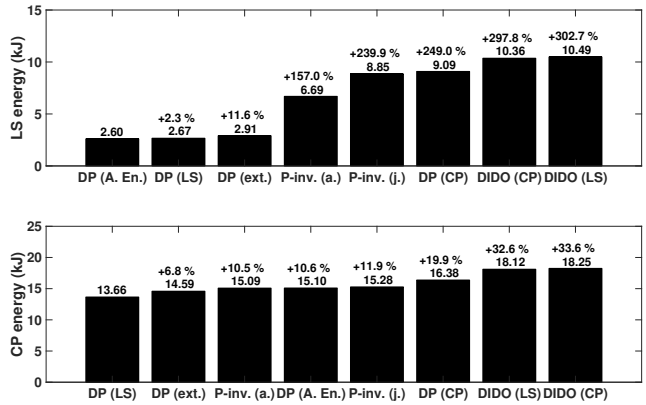


Figure A5. Vertical path energy consumption replicated using middle feasible extension cylinder length at the beginning of the path.

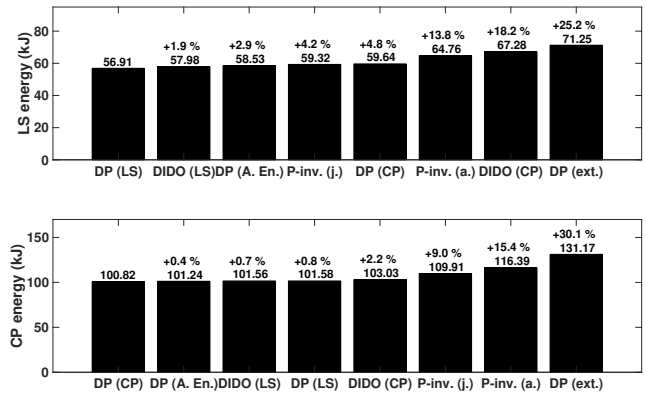


Figure A6. Half-cycle energy consumption replicated using middle feasible extension cylinder length at the beginning of the path.

References

1. Ortiz Morales, D.; Westerberg, S.; La Hera, P.X.; Mettin, U.; Freidovich, L.; Shiriaev, A.S. Increasing the level of automation in the forestry logging process with crane trajectory planning and control. *J. Field Robot.* **2014**, *31*, 343–363.
2. Westerberg, S. Semi-Automating Forestry Machines: Motion Planning, System Integration, and Human-Machine Interaction. Ph.D. Thesis, Umeå Universitet, Umeå, Sweden, 2014.
3. Whitney, D.E. Resolved motion rate control of manipulators and human prostheses. *IEEE Trans. Man Mach. Syst.* **1969**, *10*, 47–53.
4. Liegeois, A. Automatic supervisory control of the configuration and behavior of multibody mechanisms. *IEEE Trans. Syst. Man. Cybern.* **1977**, *7*, 868–871.
5. Hollerbach, J.M.; Suh, K. Redundancy resolution of manipulators through torque optimization. *IEEE J. Robot. Autom.* **1987**, *3*, 308–316.
6. Suh, K.; Hollerbach, J.M. Local versus global torque optimization of redundant manipulators. In Proceedings of the 1987 IEEE International Conference on on Robotics and Automation (ICRA), Raleigh, NC, USA, 31 March–3 April 1987; Volume. 4, pp. 619–624.
7. Chan, T.F.; Dubey, R.V. A weighted least-norm solution based scheme for avoiding joint limits for redundant joint manipulators. *IEEE Trans. Robot. Autom.* **1995**, *11*, 286–292.
8. Deo, A.; Walker, I. Minimum effort inverse kinematics for redundant manipulators. *IEEE Trans. Robot. Autom.* **1997**, *13*, 767–775.
9. Flacco, F.; de Luca, A.; Khatib, O. Control of redundant robots under hard joint constraints: Saturation in the null space. *IEEE Trans. Robot.* **2015**, *31*, 637–654.
10. Beiner, L.; Mattila, J. An improved pseudoinverse solution for redundant hydraulic manipulators. *Robotica* **1999**, *17*, 173–179.
11. Mattila, J.; Koivumaki, J.; Caldwell, D.; Semini, C. A Survey on Control of Hydraulic Robotic Manipulators with Projection to Future Trends. *IEEE/ASME Trans. Mechatron.* **2017**, *31*, 637–654.
12. Mettin, U.; Westerberg, S.; Shiriaev, A.S.; La Hera, P.X. Analysis of human-operated motions and trajectory replanning for kinematically redundant manipulators. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), St. Louis, MO, USA, 10–15 October 2009; pp. 795–800.
13. Löfgren, B. Kinematic Control of Redundant Knuckle Booms with Automatic Path Following Functions. Ph.D. Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 2009.
14. Vukobratovic, M.; Kircanski, M. A dynamic approach to nominal trajectory synthesis for redundant manipulators. *IEEE Trans. Syst. Man Cybern.* **1984**, *SMC-14*, 580–586.
15. Hirakawa, A.R.; Kawamura, A. Trajectory planning of redundant manipulators for minimum energy consumption without matrix inversion. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Albuquerque, NM, USA, 25 April 1997; Volume 3, pp. 2415–2420.
16. Vael, G.E.; Achten, P.A.; Fu, Z. *The Innas Hydraulic Transformer—The Key to the Hydrostatic Common Pressure Rail*; Technical Report; Society of Automotive Engineering, Inc. (SAE): Warrendale, PA, USA 2000.
17. Linjama, M.; Vihtanen, H.; Sipola, A.; Vilenius, M. Secondary controlled multi-chamber hydraulic cylinder. In Proceedings of the 11th Scandinavian International Conference on Fluid Power (SICFP), Linköping, Sweden, 2–4 June 2009.
18. Xue, Y.; Yang, J.; Shang, J.; Xie, H. Design and optimization of a new kind of hydraulic cylinder for mobile robots. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2015**, doi:10.1177/0954406215570106.
19. Choi, M.H. Redundancy resolution by minimization of joint disturbance torque for independent joint controlled manipulators. In Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Atlanta, GA, USA, 19–23 September 1999; pp. 392–397.
20. Flacco, F.; De Luca, A.; Khatib, O. Motion control of redundant robots under joint constraints: Saturation in the null space. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Saint Paul, MN, USA, 14–18 May 2012; pp. 285–292.
21. Nurmi, J.; Mattila, J. Global energy-optimised redundancy resolution in hydraulic manipulators using dynamic programming. *Autom. Constr.* **2017**, *73*, 120–134.
22. Ross, I.M.; Karpenko, M. A review of pseudospectral optimal control: From theory to flight. *Annu. Rev. Control* **2012**, *36*, 182–197.

23. Krus, P. On Load Sensing Fluid Power Systems: With Special Reference to Dynamic Properties and Control Aspects. Ph.D. Thesis, Linköping University, Linköping, Sweden, 1988.
24. Wu, D.; Burton, R.; Schoenau, G.; Bitner, D. Establishing operating points for a linearized model of a load sensing system. *Int. J. Fluid Power* **2002**, *3*, 47–54.
25. Jayaraman, G.P.; Lunzman, S.V. Modeling and analysis of an electronic load sensing pump. In Proceedings of the IEEE International Conference on Control Applications (CCA), Denver, CO, USA, 28–30 September 2011; pp. 82–87.
26. BR-Pump-Data. Available online: https://tinyurl.com/#kzbbksj (accessed on 6 May 2017).
27. Sciavicco, L.; Siciliano, B. *Modelling and control of robot manipulators*; Springer Science & Business Media: London, UK, 2000.
28. Armstrong-Hélouvry, B.; Dupont, P.; de Wit, C.C. A survey of models, analysis tools and compensation methods for the control of machines with friction. *Automatica* **1994**, *30*, 1083–1138.
29. Merritt, H.E. *Hydraulic control Systems*; John Wiley & Sons: New York, NY, USA, 1967.
30. Aranovskiy, S.; Losenkov, A.; Vazquez, C. Position control of an industrial hydraulic system with a pressure compensator. In Proceedings of the 22nd Mediterranean Conference of Control and Automation (MED), Palermo, Italy, 16–19 June 2014; pp. 1329–1334.
31. Kirk, D.E. *Optimal Control Theory: An Introduction*; Courier Dover Publications: Mineola, NY, USA, 2012.
32. Sundstrom, O.; Guzzella, L. A generic dynamic programming Matlab function. In Proceedings of the IEEE Control Applications (CCA) & Intelligent Control (ISIC), St. Petersburg, Russia, 8–10 July 2009; pp. 1625–1630.
33. Böhme, T.J.; Frank, B. Dynamic programming. In *Hybrid Systems, Optimal Control and Hybrid Vehicles*; Springer: Cham, Switzerland, 2017; pp. 199–214.
34. van Berkel, K.; de Jager, B.; Hofman, T.; Steinbuch, M. Implementation of dynamic programming for optimal control problems with continuous states. *IEEE Trans. Control Syst. Technol.* **2015**, *23*, 1172–1179.
35. Shin, K.G.; McKay, N.D. A dynamic programming approach to trajectory planning of robotic manipulators. *IEEE Trans. Autom. Control* **1986**, *31*, 491–500.
36. Jazar, R.N. *Theory of Applied Robotics: Kinematics, Dynamics, and Control*; Springer: New York, NY, USA, 2010.
37. Biagiotti, L.; Melchiorri, C. *Trajectory Planning for Automatic Machines and Robots*; Springer Science & Business Media: Berlin, Germany, 2008.
38. ISO-9283. Manipulating Industrial Robots – Performance Criteria and Related Test Methods, 1998.
39. Fortgang, J.D.; George, L.E.; Book, W.J. Practical implementation of a dead zone inverse on a hydraulic wrist. In Proceedings of the ASME International Mechanical Engineering Congress and Exposition (IMECE), New Orleans, LA, USA, 17–22 November 2002; pp. 149–155.
40. Vazquez, C.; Aranovskiy, S.; Freidovich, L.; Fridman, L. Second order sliding mode control of a mobile hydraulic crane. In Proceedings of the IEEE 53rd Annual Conference on Decision and Control (CDC), Los Angeles, CA, USA, 15–17 December 2014; pp. 5530–5535.
41. Bu, F.; Yao, B. Nonlinear adaptive robust control of hydraulic actuators regulated by proportional directional control valves with deadband and nonlinear flow gains. In Proceedings of the IEEE American Control Conference, Chicago, IL, USA, 28–30 June 2000; Volume 6, pp. 4129–4133.
42. Koivumäki, J.; Mattila, J. High performance nonlinear motion/force controller design for redundant hydraulic construction crane automation. *Autom. Constr.* **2015**, *51*, 59–77.
43. Nielsen, B.K. Controller Development for a Separate Meter-in Separate Meter-Out Fluid Power Valve for Mobile Applications. Ph.D. Thesis, Institut for Energiteknik, Aalborg Universitet, Aalborg, Denmark, 2005.
44. Pedersen, H.C.; Andersen, T.O.; Nielsen, B.K. Comparison of Methods for Modeling a Hydraulic Loader Crane With Flexible Translational Links. *J. Dyn. Syst. Meas. Control* **2015**, *137*, 101012.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-3974-9

ISSN 1459-2045