A Deformation Analysis Method for the Metrological ATLAS Cavern Network at CERN

- LB
- CERN

Angelika Lippitsch

Dissertation

Graz University of Technology

#22505456

Thesis - 2007 - Lippitsch

Graz, May 2007

Series Editor

Fritz K. Brunner

Engineering Geodesy and Measurement Systems
Graz University of Technology
Steyrergasse 30, A-8010 Graz, Austria
http://www.igms.tugraz.at

Previously published:

H. Hartinger, Development of a Continuous Deformation Monitoring System using GPS, 2001

A. Wieser, Robust and fuzzy techniques for parameter estimation and quality assessment in GPS, 2002

E. Grillmayer, Untersuchungen systematischer Fehlereinflüsse bei Messungen mit dem Kreisel DMT Gyromat 2000, 2002

H. Woschitz, System Calibration of Digital Levels: Calibration Facility, Procedures and Results, 2003

W. Lienhart, Analysis of Inhomogeneous Structural Monitoring Data, 2007

Engineering Geodesy – TU Graz

Angelika Lippitsch

A Deformation Analysis Method for the Metrological ATLAS Cavern Network at CERN

Shaker Verlag
Aachen 2007

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: Graz, Techn. Univ., Diss., 2006

The dissertation was submitted to the Faculty of Mathematical and Physical Sciences at the Graz University of Technology for achieving the academic degree of Doctorate of Technical Sciences.

Examination Committee:

Examiner: Univ.-Prof. Dr. Fritz K. Brunner
Second examiner: Univ.-Prof. Dr.-Ing. Otto Heunecke

Day of oral examination: December 20, 2006

Copyright Shaker Verlag 2007
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8322-6289-1
ISSN 1864-2462

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

Abstract

The Large Hadron Collider (LHC) project at CERN includes the construction of four large physics experiments, which will study particle collisions. Each particle detector needs to be precisely aligned with respect to the accelerator beam line by survey measurements. One of these particle detecting experiments is called ATLAS.

Configurations of the geodetic underground cavern network are constrained due to access and space limitations. Sighting restrictions increase as the installation of detector parts in the cavern progresses. Consequently, the reliability of the network reduces. Additionally, deformations of the cavern as well as access structures affect the geodetic network and need to be considered. Dedicated network measurements can only be carried out on an irregular and sparse basis. For processing the inhomogenous and hybrid measurements an adaptive Kalman Filter (KF) is developed interpreting the cavern network as a kinematic system. This enables to handle changing network configurations easily, as well as maintaining a higher level of reliability in the network compared to individually adjusted network measurements. With such an algorithm it is possible to evaluate survey measurements more efficiently, giving accurate estimations for the point positions and corresponding

error information at times of measurements but also to predict future positions and error estimates.

The application of the developed method to the processing of simulated and real survey data sets for the ATLAS cavern network demonstrates the algorithm's superior performance and advantageous capabilities compared to conventional processing methods.

Zusammenfassung

Im Rahmen des Large Hadron Collider (LHC) Projektes am CERN werden vier große Teilchendetektoren Kollisionen untersuchen. Jeder dieser Detektoren muss mit Hilfe von geodätischen Messungen in Bezug auf den Teilchenstrahl ausgerichtet werden. Eine dieser Detektoranlagen ist ATLAS.

Zahlreiche Einschränkungen für die Konfiguration von geodätischen Netzen ergeben sich aus dem Umfeld einer Kaverne. Behinderungen durch bauliche Einrichtungen und Detektorbauteile, die mit zunehmendem Baufortschritt in der Kaverne Platz finden, nehmen mit der Zeit zu. Die Möglichkeiten für geodätische Messungen werden immer mehr eingeschränkt und die Zuverlässigkeit im geodätischen Netz wird dadurch reduziert. Das geodätische Netz wird zusätzlich von Deformationen der Kavernenstruktur und Plattformen beeinflusst. Spezielle Netzwerkmessungen können nur selten und in unregelmässigen Abständen durchgeführt werden.

Ein adaptiver Kalman Filter (KF) wird entwickelt, der die kinematische Interpretation eines Netzwerkes erlaubt. Veränderliche Netzwerkkonfigurationen können im KF Formulismus einfach behandelt werden. Die Zuverlässigkeit des Netzwerkes kann auf einem hohen Niveau ge-

halten werden. Messdaten werden im KF Algorithmus effizienter analysiert, als im Ausgleich von einzelnen Messepochen und ergeben präzisere und zuverlässigere Koordinatenergebnisse und Genauigkeitsinformationen. Für Epochen ohne Messdaten können Koordinaten und entsprechende Genauigkeitsinformationen prädiziert werden.

Die Anwendung der entwickelten Methode auf simulierte und reale Vermessungsdaten für das ATLAS Kavernennetzwerk zeigt die hohe Leistungsfähigkeit des Algorithmus.

Acknowledgments

I would like to thank Prof. Fritz K. Brunner, head of the Institute of Engineering Geodesy and Measurement Systems at Graz University of Technology for supervising this thesis.

Many thanks go to Christian Lasseur, head of the Experimental Metrology Section at CERN, who proposed the topic of this thesis and who gave me as much time as possible to work on and finish this thesis even though the experiment installation schedules were pressing on the survey team. A big 'Merci!' to all members of CERN's survey group who supported me in my thesis work and in the practical surveying work during my time at CERN.

I would like to thank especially Dirk Mergelkuhl and Antje Behrens for proof reading this thesis and Vasco Correia Carmona Pires and Andreas Herty for their help in taking the photographs of CERN survey equipment.

The support and encouragement of Werner Riegler, coordinator of the Austrian Doctoral Student Program, and of Christian Joram and Lucie Linssen of CERN's Physics department have been very important enabling me to finish this thesis.

I am very grateful to have Alexander Koschik as my partner at my side. His help and support for me to work on and write this thesis and his dedication to understand surveying analysis techniques, often strange to a physicist's understanding, has been unrivalled.

Contents

Abstract i

Zusammenfassung iii

Acknowledgements v

Contents vii

1 Introduction 1
- 1.1 CERN – LHC 1
- 1.2 Motivation–ATLAS: Problem description 4
- 1.3 Scope of the thesis 8

2 CERN ATLAS cavern network 11
- 2.1 Coordinate systems for LHCPoint 1 – ATLAS 12
- 2.2 The ATLAS cavern 13
  - 2.2.1 Underground structures 13
  - 2.2.2 Geological and geotechnical aspects 15
  - 2.2.3 Access structures 20

2.3 The ATLAS detector 23
2.3.1 Detector systems 23
2.3.2 Installation schedule 26
2.3.3 Detector access scenarios 29
2.3.4 Positioning requirements 30
2.4 The metrological ATLAS cavern network 32

3 Metrology at CERN applied to detector positioning 37
3.1 CERN reference systems 38
3.2 Network point monumentation 41
3.3 Measurement methods 50

4 Standard network and deformation analysis 59
4.1 Network analysis 59
4.1.1 The geodetic datum 60
4.1.2 Accuracy 62
4.1.3 Reliability 63
4.2 Standard deformation analysis 67
4.3 Conclusion 72

5 Alternative adjustment of deforming networks 73
5.1 Deformation models 74
5.2 Kalman Filter in geodetic applications 75
5.3 Kalman Filter basics 76
5.3.1 State–space model 77
5.3.2 Discrete Kalman Filter 79
5.3.3 Nonlinear filtering 82
5.3.4 Innovation Analysis - Adaptive filtering 85
5.3.5 Some characteristics of the KF 88
5.4 KF equations for a kinematic 3D geodetic network . . . 90
5.4.1 From equations of motion to system equations . 91
5.5 Adaptive Kalman Filter for a kinematic network 96
5.5.1 Summary of KF terms for kinematic 3D network 97
5.5.2 Innovation analysis 100
5.5.3 Stochastic stabilization 107
5.5.4 KF characteristic terms 110
5.5.5 Gain matrix **K** 112
5.5.6 Selection of system model – Multiple-model filter 114
5.6 Including additional information in KF 115

5.6.1 Formulating conditions as additional information 115
5.7 Conclusion 117

6 Application and results of adaptive Kalman Filter 119
6.1 Implementation of the adaptive KF algorithm 120
6.2 Small simulated network data 124
6.2.1 Network configuration and observations 127
6.2.2 Analysis and comparison of adaptive KF results 128
6.2.3 Influence of stochastic stabilization on gain matrix 144
6.2.4 Including additional information in adaptive KF 154
6.2.5 Conclusion for the small simulated network . . . 156
6.3 Simulated ATLAS cavern network data 157
6.3.1 Network data simulation 157
6.3.2 Network analysis 162
6.3.3 Analysis and comparison of adaptive KF results 165
6.3.4 Conclusion for simulated ATLAS cavern network 175
6.4 Real ATLAS cavern network data 182
6.4.1 The real ATLAS cavern network 182
6.4.2 ATLAS cavern network measurements 183
6.4.3 Analysis and comparison of adaptive KF results 189
6.4.4 Conclusion for the real ATLAS cavern network . 209
6.5 Conclusions for application of the adaptive KF 211

7 Conclusions 213

Appendix 217

A S-Transformation 217

B Structure of the gain matrix K 219

List of References 223

Curriculum vitae 231

1

Introduction

1.1 CERN – LHC

The European Organization for Nuclear Research (CERN – Centre Européen pour la Recherche Nucléaire) is currently concentrating its efforts on the new project LHC (Large Hadron Collider) which will be the most powerful particle accelerator in the world.

Particle accelerators are used to generate high energy beams of either protons or electrons which are then collided at specific locations creating new particles which allow to study the structure of matter. Particle detectors are capable of measuring energy, mass and charge of these newly created particles. In order to create and study particles with higher momentum, the colliding particle beams have to be of higher energy which has driven accelerator technology in recent decades to build accelerators with ever higher energy. This development will reach its momentary maximum with LHC which will accelerate two counter-rotating proton beams to 7 TeV per proton. In modern high-energy physics experiments, large multi-layered detectors surround the particle beam collision point, the so called *interaction point* (IP). Each layer of the detector (grouped into individual detector subsystems) serves a separate function in tracking and identifying each of the many particles that may be produced in a single collision.

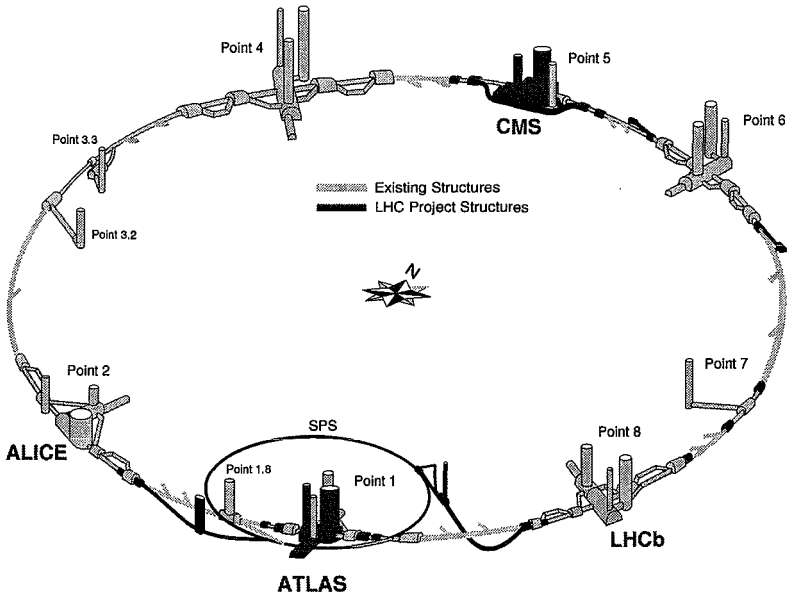


Figure 1.1: *The LHC project. Existing and newly constructed underground openings including the new LHC experiments ATLAS, ALICE, CMS and LHCb,* ©CERN.

The LHC machine will accelerate protons along an almost circular trajectory of approximately 27 km. It is installed in an existing tunnel which accommodated the LEP (Large Electron Positron) Collider until its shut-down in November 2000. The existing tunnel and shafts could be largely adopted but some additional excavation work was necessary. The LHC tunnel is situated on the French – Swiss border in the Leman basin, $50 - 175$ m underground.

Part of the LHC project are four high energy physics experiments which have been constructed and installed on four beam collision points along the LHC particle beam line, as can be seen in Figure 1.1.

All LHC experimental detectors consist of multiple subsystems. Each of these systems tests for a special set of properties of particles which are created in the beam collisions. The trajectories of the decay particles are traced when passing through the individual detector systems.

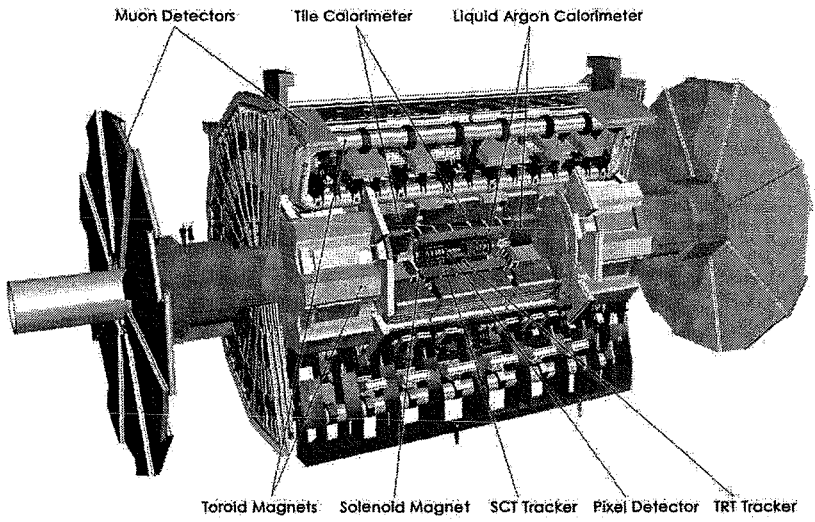


Figure 1.2: *Model of the ATLAS detector: Detector systems. The size of the detector is compared to the size of persons in the illustration, ATLAS Experiment image,* © CERN, [3].

Since the detectors are inside a powerful magnetic field, the curvature of the observed particle trajectories are a function of their momentum and charge. The momentum gives the mass of the observed particle and together with the energy measured by the calorimeters the particle can be identified. To get correct and precise results it is absolutely essential that the positions of the detectors and all individual parts are known with respect to the IP and accelerator beam line and also relative to each other within one experiment. Accuracy requirements in some systems can be as small as 30 μm. Compared to the size of the detectors and considering their complex structures, these requirements are very challenging for both manufacturing, installation, positioning, geometrical control and analysis.

The two biggest LHC experiments are ATLAS (A Toroidal LHC ApparatuS) at Point 1 and CMS (Compact Muon Solenoid) at Point 5. The two corresponding underground caverns UX15 at Point 1 and UX5 at Point 5 had to be newly excavated whereas caverns at Point 2 and Point 8 could be 'recycled' from former LEP experiments.

The ATLAS detector is about the size of a five story building and is placed in the UX15 cavern having a dimension of approximately $53 \times 30 \times 35$ m. The detector is comprised of the following major systems, see also Figure 1.2: Inner Tracking Detector (including Pixel Detector, SCT Tracker and TRT Tracker), Solenoidal Magnet, Electromagnetic Calorimeter (Liquid Argon Calorimeter), Hadron Calorimeter (Tile Calorimeter) and Muon Detectors (including Toroid Magnets). Each of these components are again assembled of numerous parts. Alignment and position of each detector part has to be established and controlled during the manufacturing and installation process and also later under operation. Different requirements apply for different detector systems.

Once accelerator and detectors will be in operation additional alignment systems installed inside individual detector systems will provide online information of the relative geometry of related detector parts. These systems will yield online information which will be used to correct the data gathered by the detectors at the collision events. As these systems are set up to provide only small relative corrections they have to rely on precise a priori positioning and constantly monitored alignment with respect to the beam line. They do not make the precise positioning obsolete but rather depend on it. These online relative detector alignment systems are not discussed in this thesis.

Excavation work for the new LHC experiments started in 1998. First collisions of particles accelerated by the LHC are expected by the end of 2007. Approximately 3000 persons form the scientific and technical staff at CERN. Each year more than 5000 researchers representing collaborating institutes come to CERN. The ATLAS collaboration includes 163 institutes from 34 countries.

For more information about CERN and the LHC project visit `http://www.cern.ch`.

1.2 Motivation–ATLAS: Problem description

Installation of the new large physics experiments is a very critical task in the LHC project. The accuracy requirements in some locations are very demanding and often call for the application and development of special techniques. Also logistics behind such a project are unique and

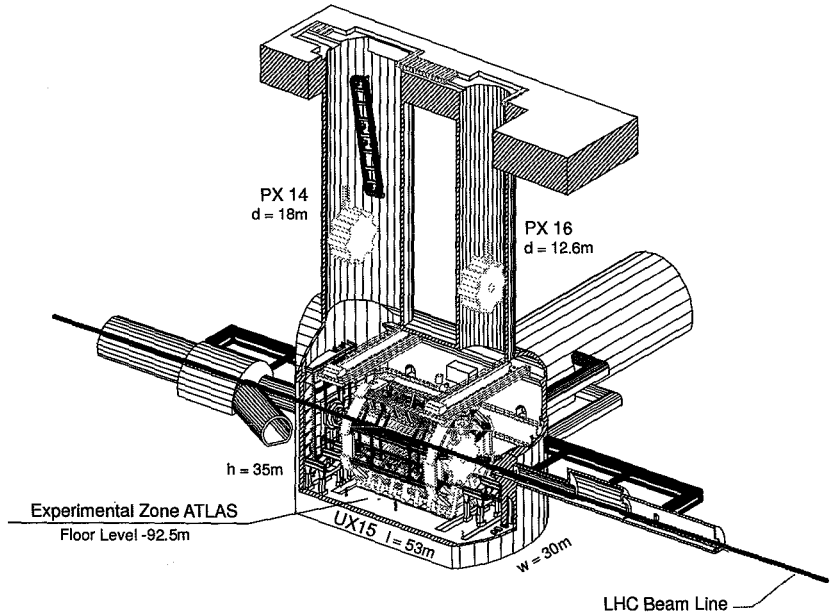


Figure 1.3: *Lowering of detector components into the cavern, [2].*

require long planning and preparation time before carrying out the actual tasks. Whenever possible pre-assembly steps are carried out at other locations than the actual experimental zones which represent very confined environments. In case of the biggest experiment, ATLAS, pre-assembly is only possible to a relatively small extend, simply because of the detector's design. ATLAS detector systems are very complex and their structure reminds of a 'Russian doll' where one part is encased in another, not accessible from the outside. Thus the major part of the ATLAS assembly process has to be carried out in the underground cavern, see Figure 1.3. Space and access possibilities are very limited in such a confined environment and make the surveyor's work in many aspects challenging. See also Figure 1.4 and Figure 1.5 which show some snapshots of the installation progress in the ATLAS cavern.

The installation of detector parts in the cavern is guided and controlled by geodetic measurements to enable and ensure precise and ac-

curate positioning. These measurements refer to a local cavern network, which is linked to the LHC machine geometry representing the theoretical accelerator beam line. A network had been designed, considering mechanical and geometrical constraints and limitations in order to obtain a sufficiently reliable network. This layout is referred in the following as the *nominal network*. It has been used as guidance for installation of network points in the real ATLAS cavern. The network is realized by geodetic monuments which are either attached to the cavern walls, mounted on metal structures, which serve as support for various infrastructure installations and as access for the personal or embedded in the cavern floor.

As the ATLAS cavern in particular has been newly excavated in its major extend, it has to be expected that the cavern will both deform and possibly move with respect to the tunnel housing the accelerator. This cavern deformation directly affects network points along the cavern walls and those embedded in the cavern floor. Additionally it translates to movements and/or deformations of the metal support and access structures where many other network point monuments are mounted. These structures are also expected to deform independently as loads are applied intentionally or accidentally. As the network point monuments are located in a working environment they are protected from external influences as much as possible, but still might be subject to unexpected forces. Possible effects of such events have to be considered and controlled. This complex deformation process directly affects the geodetic network, its internal geometry and also reliability. As the cavern network is defined with respect to the nominal accelerator beam line, possible deformations or movements make this relationship very important and would ask for a very strong geometrical connection. This is unfortunately impossible due to vigorous limitations on the network design. The network thus needs to be analyzed very carefully for any possible deformation in order to fulfill the demands over the whole installation and operational period of the experiment.

In the early stage of the installation process the cavern is mostly empty, see for example Figure 1.4. As soon as the first network points are mounted in the cavern a first small network is measured. At that time it is still possible to measure this network in a very good configuration i.e. many measurements even across the later filled detector space, as the sighting limitations are not very strong. In parallel with

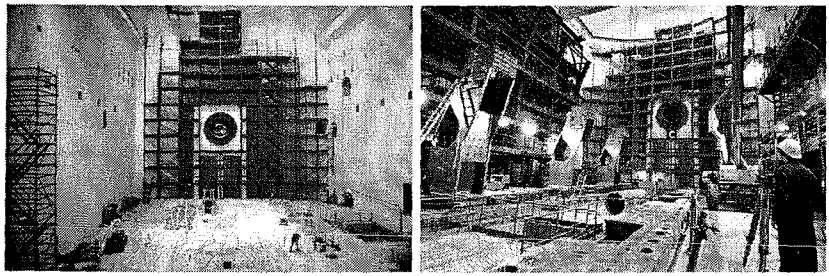


Figure 1.4: *Photographs of the ATLAS cavern in August 2003 (left): beginning of infrastructure and access structure installation shortly after the delivery of the cavern to CERN, and November 2003 (right): during installation of ATLAS feet,* ©CERN.

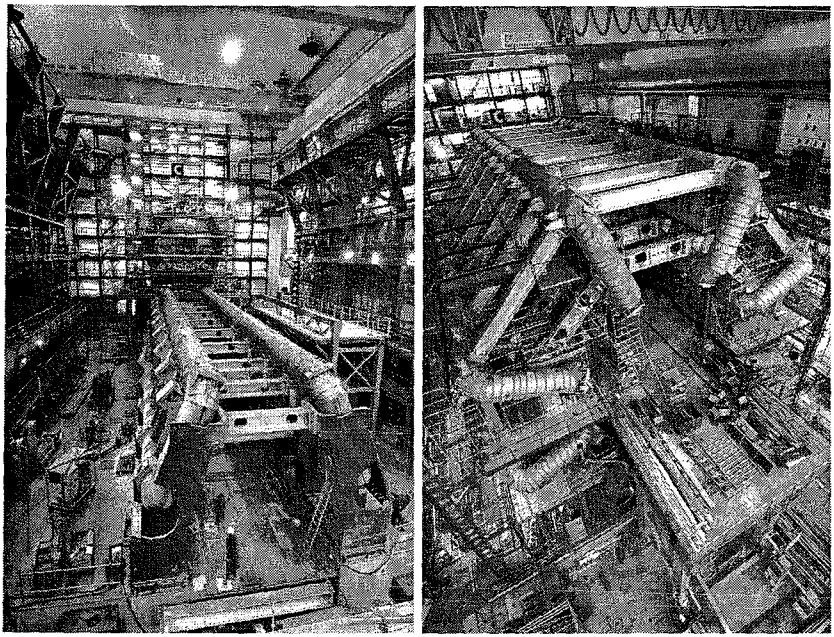


Figure 1.5: *Photographs of the ATLAS cavern in December 2004 (left) at the beginning of installation of the Barrel Toroid, 2 coils already installed. Tile calorimeter is assembled on cavern side C. Right: September 2005 with all 8 Barrel Toroid coils installed and start of Muon Chamber insertion into this structure,* ©CERN.

the progressing infrastructure installation, more network points are installed until the full network is in place. But as more infrastructure and detector parts are installed, sightings between the network points are increasingly limited, see Figure 1.5. This also affects the network as it degrades the configuration and thus its quality in terms of control and reliability. Having a stable network with reliable reference information is essential in this situation. Careful measurement planning in the cavern network is important in difficult configurations and only together with thorough statistical analysis of the survey data a high level of reliability in the network can be maintained.

Thus, for the survey work to support detector installation two major problems have to be faced:

- Deformation of the cavern network directly affecting its stability and accuracy.
- Degrading network configuration due to increasing geometrical limitations with progressing detector and infrastructure installation resulting in poorer network reliability.

In spite of these two problems it is necessary to maintain a high level of reliability in the network to ensure accurate guidance of installation and positional control of detector parts.

Several known techniques offer themselves for application in order to deal with both problems individually. Here, however, the problems are closely related, hence they cannot be addressed individually. An alternative solution has been sought to treat these two problems in a combined way and is described in this thesis providing a practical solution.

1.3 Scope of the thesis

A metrological network in a large physics experiment cavern like ATLAS represents a special situation for an engineering surveying application. Therefore more details are given about this special situation by means of problems and demands that have to be handled before a solution to cope with this special situation is presented. Thus this thesis is structured in the following way:

- In Chapter 2 information on the ATLAS cavern, the ATLAS detector and the metrological ATLAS cavern network is presented and discussed. This includes information about the geological and geotechnical situation at LHC Point 1, some details about the excavation and construction works and also details about infrastructure and access structures in the cavern and this information's relevance to the study of cavern and network deformation. The metrological network is discussed considering the nominal network design.

- In Chapter 3 a short summary of metrological measurement techniques employed by CERN's survey group is given, focusing on techniques applied in the ATLAS cavern network and its link to LHC reference geometry. Special attention is paid to specialized survey systems. This chapter also includes a section on network point monumentations commonly used at CERN and particularly employed in the ATLAS cavern network and vicinity.

- In Chapter 4 standard network and deformation analysis techniques are shortly discussed considering their suitability and limitations for the task at hand.

- In Chapter 5 an alternative approach to network management and deformation analysis is presented. The proposed method is the application of the Kalman Filter technique to a 3D kinematic network. With this approach it becomes possible to treat problems of a deforming network in degrading configuration in a very efficient way.

- In Chapter 6 results of the algorithm introduced in Chapter 5 are presented for both simulated and real measurement data. Data sets are simulated for two examples: One being a very small and simple network which serves as an example to present the algorithm's performance in a very straightforward and direct way. Secondly data for the theoretical ATLAS cavern network has been simulated considering early installation schedules and deformation scenarios based both on assumptions and information discussed in Chapter 2. With this simulation example special features of the ATLAS cavern network can be investigated and critical issues in

the network or measurement configurations identified. Finally, results for real ATLAS cavern network data for a period of three years is analyzed and presented showing the capability of the proposed algorithm to deal with less favorable data.

- A summary in Chapter 7 concludes this thesis.

2

CERN ATLAS cavern network

In this chapter basic engineering information about the ATLAS experimental zone (i.e. cavern and surrounding underground structures), the ATLAS detector and the metrological cavern network are summarized. It is discussed whether it is possible to derive a deformation model from information about expected forces acting on the underground complex or not. Such a dynamic model could be included in the deformation analysis of survey data gathered in the cavern network. The ATLAS cavern network is discussed in its nominal design layout. First, however, a short overview of the different coordinate systems in use in the ATLAS cavern is given.

2.1 Coordinate systems for LHC Point 1 – ATLAS

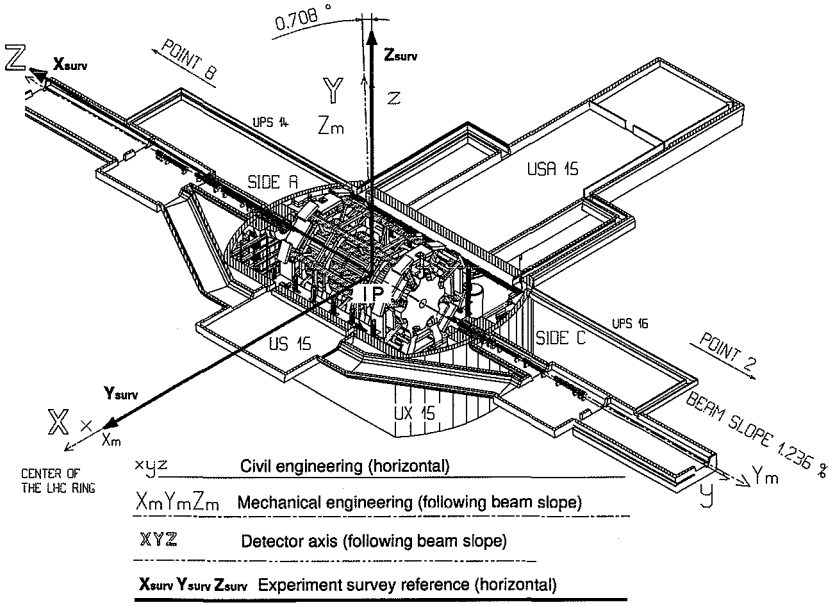


Figure 2.1: *Coordinate systems in the ATLAS cavern, [2]*

The origin of all coordinate systems defined for the ATLAS cavern or the ATLAS detector is the theoretical interaction point IP, [2]. For an overview see Figure 2.1, in which the four ATLAS coordinate systems are shown.

Detector axis system The X-axis is horizontal, pointing towards the center of the LHC ring. The Y-axis is perpendicular to the X-axis and to the beam axis, positive upwards. It is inclined by 1.236% (0.01236 rad) or 0.7082° with respect to the local vertical. The Z-axis is aligned with the beam direction, to create a right-handed Cartesian coordinate system.

Mechanical engineering system In this system the X-axis is horizontal and pointing towards the center of the LHC ring. The Y-axis follows the beam slope upwards. The Z-axis completes a left-handed Cartesian coordinate system. It is inclined by 0.01236 rad with respect to the local vertical. This system is used by mechanical engineering groups.

Civil engineering system The X-axis is horizontal and pointing towards the center of the LHC ring. The Z-axis is perpendicular to the X-axis and points upwards along the local vertical. The Y-axis completes the Cartesian coordinate system. It is horizontal and is therefore inclined by 0.01236 rad with respect to the beam direction. This system is used by the civil engineering group.

Experiment survey reference system The ATLAS survey coordinate system is a local, horizontal coordinate system. The Y-axis is horizontal and points towards LHC center. The X-axis is horizontal pointing towards Point 8. The Z-axis is following the local vertical, positive upwards and thus completes the right-handed Cartesian coordinate system. This system is used by the experiment survey group.

2.2 The ATLAS cavern

2.2.1 Underground structures

The ATLAS cavern is situated at CERN LHC Point 1, on Swiss territory. The cavern floor is approximately 92 m below ground. The cavern itself has the internal dimensions of length= 53 m, width= 30 m and height= 35 m. For illustration see Figure 1.1 and Figure 1.3.

The cavern is connected to other underground openings in this area, thus forming a complex of underground openings (see Figure 2.2). These other openings include transport and security shafts, transfer tunnels, caverns, chambers and galleries for service and infrastructure installations.

The main cavern is referred to (see Figure 2.2) as UX15, its axis is parallel to the direction of the beam line but it is horizontal. The large cavern USA15 is perpendicular to UX15 and holds, among others, cryogenic services. It has a diameter of 20 m and a length of 62 m. It

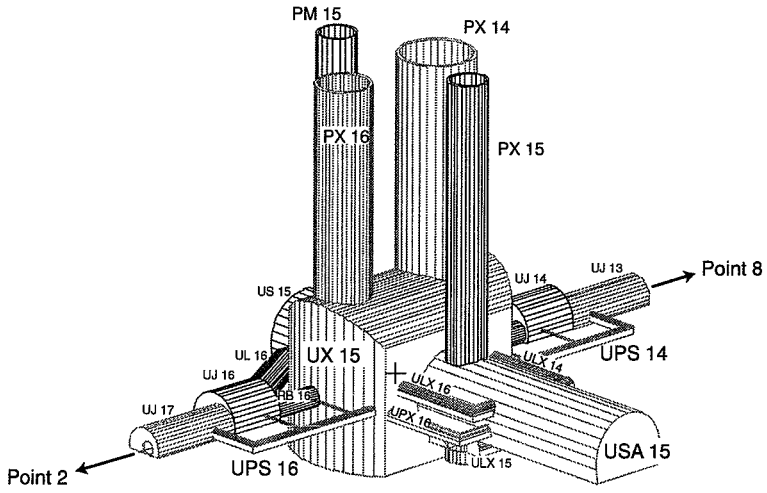


Figure 2.2: *ATLAS: The complex of underground structures at LHC Point 1, [2].*

is accessed by shaft PX15, which was built for the LEP installation and which is now the main personnel access to the new underground structures.

The new circular shafts PX14 and PX16, with internal diameters of 18 m and 12.6 m respectively, provide material access to the main cavern UX15. Other existing structures include the service shaft PM15 which is used to access the LHC tunnel and the service cavern US15. The new underground structures were built around and into existing structures. The galleries UPS14 and UPS16 serve as service galleries for metrological measurements (*survey galleries*) and special measurements systems, which are described in Chapter 3.

The LHC beam is inclined down from Point 2 to Point 8. This inclination has to be taken into account in the detector installation process as the cavern floor is horizontal. The right slope angle of the beam line is 0.01236 rad or 0.7082°. The IP of ATLAS is 11.37 m from the cavern floor, which also defines the origin of the local ATLAS coordinate system(s), see Section 2.1. The lateral cavern side towards the adjacent cavern USA15 is referred to as USA, the opposite lateral side US. The end side towards Point 8 is called side A, and the opposite side towards Point 2 side C.

2.2.2 Geological and geotechnical aspects

A short overview of the geological situation at Point 1 as well as some geotechnical aspects of the cavern construction shall be given here, based on reports by Rammer [50, 51, 52], and Guitton [17]. Although the study of these extensive reports in search for information about the expected cavern deformation behavior after construction has not turned out satisfactory, this section is still included to give a broader view of the situation. A short conclusion at the end of this section summarizes the findings for geological and geotechnical information.

The CERN site is situated in the Leman Basin, enclosed by the Alps in the south-east and the Jura mountains to the north west. The basin is filled by sedimentary deposits, collectively called molasse. These deposits comprise a complex, alternating sequence of almost horizontally bedded sandstones and marls, with a range of composite marly sandstones and sandy marls. The molasse is overlain by moraines from the glacial periods of Riss and Wurm, consisting of gravel and sands with many cobbles and boulders and varying in depth with less than 10 m at Point 1, [50].

Data on the geological and geotechnical situation stem from studies carried out before the construction of the LEP tunnel, in 1981 and 1982 and before the LHC project between 1995 and 1997. In both cases this included boreholes drilled and tested along the tunnel trajectory, including the site at Point 1. Rock samples were collected and analyzed in the laboratory and *in-situ* testing was carried out.

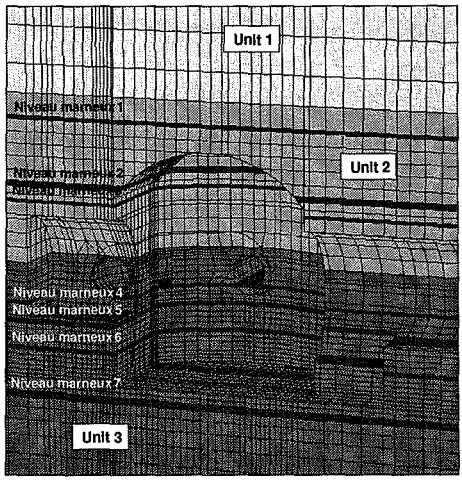


Figure 2.3: *Simplified rheological model for UX15 implemented in 3D FEM software FLAC3D version 2.00. ('Niveau marneux' stands for 'layer of marl'), [12].*

For the derivation of the geotechnical model at Point 1, it was necessary

to make some simplifications to cope with the problem considering software and computer resources available at that time. The sequence of various rock types was analyzed in detail for a selected group of representative boreholes and a simplified series of bands with similar properties derived.

As basis for subsequent modelling these bands have been simplified even more into three basic rock mass units, see Figure 2.3. The upper unit consists predominantly of sandstones with some intermediate layers of marls. It extends from the base of the moraine, at depths between 4 and 10 m, down to 50 m. The middle unit extends to depths more than 80 m and is the most diverse being composed by sandstones and marls but also transitional rock types. The lower unit consists predominantly of marls with sandstones and a minor proportion of transitional rock types.

The analysis of various laboratory tests allowed to derive strength and deformation parameters for each type of rock. Special attention was paid to the weaker marls especially in critical zones (e.g. just above the cavern vault) by implementing separate bands into the geotechnical model. This model was used in the design and implementation of the civil engineering works discussed in the following.

Civil Engineering Works

The civil engineering project had to consider additional constraints to the geological situation:

- Size of the detector to be installed in the experimental cavern.
- Large diameter access shafts are required to lower pieces of equipment into the cavern.
- Integration of new works into existing structures.
- Construction to be completed in 2002, LEP shutdown end of 2000.

Especially the last constraint on the time schedule made it necessary to start excavation 2.5 years prior to LEP shutdown. It was required to have as little effect as possible on the continuing accelerator operation.

The numerical analyses based on time schedules, geotechnical and geometrical models carried out by the contractor consortium EDF-Knight Piésold involved a multitude of 2D and 3D numerical calculations, [12]. The 3D analyses had the main objective to investigate the interaction of different underground openings. More detailed studies of

some critical sections in the model were carried out with a 2D approach. It enabled to address some more precise details of the geological model. Considering the tight LHC project time schedule the maximum possible amount of construction work had to be carried out before final LEP shutdown. All works were to cause minimum effects on the existing structures and no disruption to the operation of the LEP machine. This implied that explosives for excavation were forbidden thus mechanical excavators were chosen. The cavern vault had to be excavated and concrete lining applied before LEP was shutdown. Only then all the equipment in the tunnel could be dismantled and removed and the excavation of the cavern benches could proceed. This required for a special solution of a temporary support of the dead weight of the concrete cavern roof by suspension of 38 pre-stressed ground anchors of 225 tons capacity each. This support system had to work until the construction of the cavern base slab and walls when they took the weight of the concrete roof and the anchor cables became redundant, [51, 46]. See Figure 2.4 for illustration.

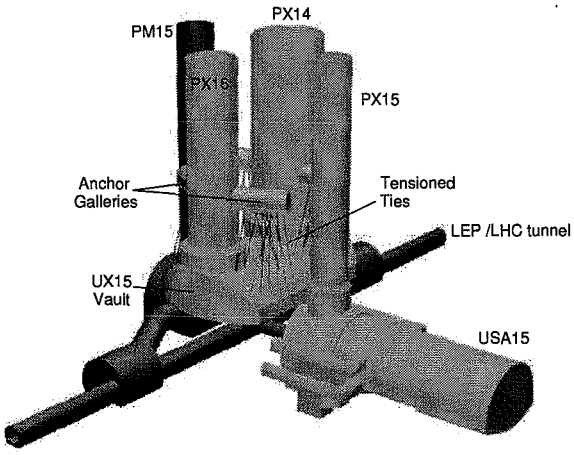


Figure 2.4: *General layout of the Point 1 underground works prior to bench excavation, [46].*

The permanent concrete lining of the cavern consists of a 5 m thick base slab, 2 m thick straight vertical side walls, 1m thick curved vertical end walls and 1.3 m thick roof arch. The primary rock support consists of fibre-reinforced shotcrete and rock bolts. To ensure watertightness an impermeable membrane is included between the shotcrete and the final concrete lining.

The zones causing most problems were shear zones at the level of the cavern vault, at the level of the LEP tunnel and also below that level, caused by layers of marly sandstone and sandy marls. This behavior was anticipated by earlier calculations (see an illustrated example in Figure 2.5) and could be observed with inclinometric measurements. The excavation procedure was adapted for these zones and thus instabilities could be sufficiently stabilized.

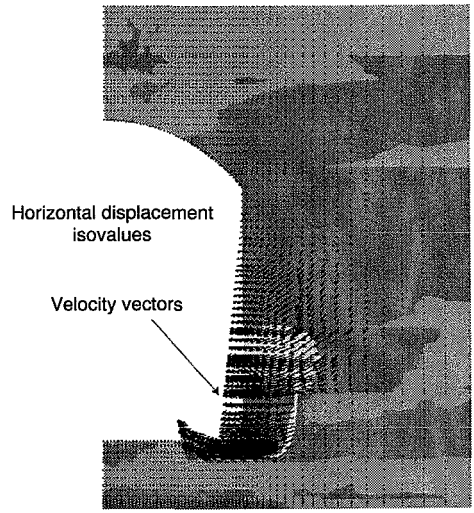


Figure 2.5: *2D calculation UX15: Isovalues horizontal displacements and velocity vectors, [12].*

Monitoring and control of the excavation work were carried out by survey and geotechnical measurements conducted by external companies. The data was used to constantly refine the numerical models. Thus also the actual load which would stress the cavern walls was determined and the final lining of the cavern was accordingly chosen. Convergence measurements which were made to optical targets on the inside of the excavated cavern were compared to predicted values of displacements. Results of such comparisons showed that the predicted values were much larger than the observed values. 3D computations based on largely simplified rheological models but including surrounding structures had predicted values 3.5 times larger than observed and had indicated an asymmetric deformation, smaller on the USA side of the cavern because of the close vicinity of the USA15 cavern which could not be observed. 2D models based on the refined rheological model but not considering any interactivity with other structures (thus symmetric deformation expected) had predicted values 5 times larger than observed. Results of these comparisons could be introduced as additional information to the deformation model. These refined models helped to finish excavation work without major problems and were used to determine the dimension of the final inner lining of the cavern.

The tolerances for the final internal dimensions of the cavern were rather tight (-0 and $+7$ cm), being softer on the positive side (thus bigger cavern). Eventually the cavern was found (by final measurements by both the contractor and CERN (laser scanner)) to be bigger by approximately $5-10$ cm.

Conclusion for cavern deformation model Geotechnical and geological information and some interesting details about the cavern construction have been summarized in this section. For the problem at hand, no quantification or model for future deformation behavior of the underground structures after completed construction could be derived. In the understanding of civil engineering, the cavern is not deforming, once the inner lining is applied, [53]. Having said that, experience from former LEP experiment underground openings had shown that remaining forces on the lateral cavern walls cause them to be pushed slightly towards the center of the cavern (approximately 1 mm per year, [29]). Additionally it is expected that the cavern floor will subside by 2 mm directly after completion of the cavern before the detector parts are installed, and extra 5.5 mm over the whole installation period caused by the load of about 7000 t of the ATLAS detector plus 1000 t of access structures and infrastructure installations. Counteracting this vertical sag, a heave of cavern floors has been observed in former LEP caverns caused by hydrostatic pressure (estimated for UX15 1.2 mm per year). Stability (defined as movements lower than 1 mm per year) is not expected to be achieved even after 15 years, [53, 7].

These values and assumptions cannot be quantified precisely and no combined deformation model is available. Considering deformation analysis in combination with survey data gathered in the cavern network, it is unfortunately not possible to include an analytical model of the cavern dynamics. However, the empirical information discussed above is used in a simulation of possible deformation scenarios of the ATLAS cavern, see Chapter 6.3. Analysis of real ATLAS cavern network data shows if the anticipated displacements and deformations takes place or not, see results in Chapter 6.4.

2.2.3 Access structures

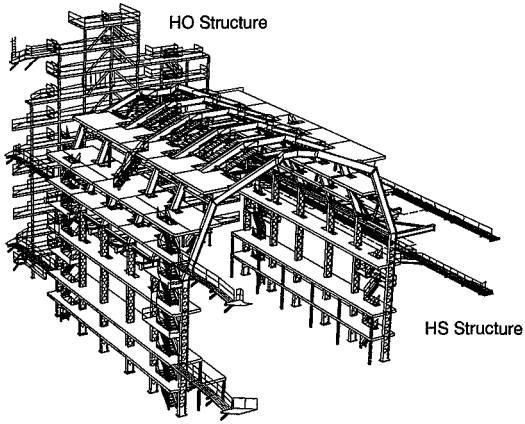


Figure 2.6: *Infrastructure and access structures HS and HO in the ATLAS cavern, [2].*

A large number of network points are mounted on access structures also serving for infrastructure installations, surrounding the detector in the cavern, see Figure 2.6.

Deformation models of these structures (especially any interconnections with a cavern deformation model and the evolvement over time) would be of interest as they could be included in a dynamic deformation analysis of survey data in the ATLAS cavern network. Models for both *HS* and *HO* structures are discussed in this section. HS structures are located along the lateral cavern walls on both US and USA side of the cavern. HO structures are erected close to both curved end walls of the cavern, on side A and side C. They are produced of structural steel and are attached to both the cavern floor and walls. There is no direct connection between HS and HO structures. For both types of structures Finite Element Models (FEM) are available, but these models have been calculated independently of each other, in different software formats and following different standards. At the end of this section conclusions are drawn for both HS and HO structure information together.

HS structure

HS structures along the lateral sides of the cavern differ in their dimension, resulting from the detector interaction point (IP) not coinciding with the center of the cavern. As there is more free space on the USA side, various infrastructure facilities have been installed on this side and thus the access structure expands further from the cavern wall.

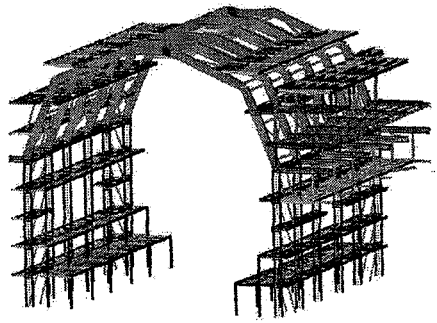


Figure 2.7: Visualization of the finite element model of the HS structure, [11].

The HS FEM model is implemented in the *Robot Finite element package*, [54]. The aim of the FEM calculations carried out at CERN were to verify the structure according to the regulations in force, French Standard EUROCODE3 and PS92 (norme parasismique française), [11]. The HS structure consists of twelve posts of profiled steel Fe360, joined in pairs by an arch section on the top plus horizontal elements of different profile types. The FEM representing the structure is made up of 3660 elements and 2370 nodes. Load cases to be considered include several groups, classified in their nature: Self weight (permanent), detector and infrastructure parts weights (permanent/variable), cryogenic equipment (permanent), thermal load (variable), magnetic load (variable), seismic load (accidental), shock on structure (accidental) and personnel mass (variable). The calculations according to EUROCODE3 define certain load combinations according to the limit states, i.e. Ultimate Limit States (ULS) and Serviceability Limit States (SLS) defined by different factors on each permanent and variable states, considering accidental loads or not. According to PS92 the loads are translated into effects on individual coordinate directions. Certain combinations of this directional effects are then evaluated.

Boundary conditions used in the calculations include connections of individual structure parts to the feet bases on the cavern floor and in cavern walls, which were assumed infinitely rigid.

Results of the FEM calculations showed that the structure fulfilled all regulations. For SLS a displacement maximum was found in vertical direction of 25 mm (downwards) on top arch beams connecting vertical posts. On average displacements were in the range of $2 - 3$ mm, dominantly in vertical direction. It is noted in the analysis report [11] that no lateral displacements of the cavern walls were taken into account in the final calculation as this posed a high level of stress on the struc-

ture. It is recommended to use connections decoupling any transversal movement of the walls. Whether such measures become apparent in the results is presented in Section 6.4.

HO structure

HO structures along the curved end walls of the ATLAS cavern differ only slightly for both facing sides, they are otherwise symmetric. The HO FEM is implemented in *ANSYS* software, Version 5.6.

The main objective of the calculation at hand [20], is the validation of the structure to carry the weight of the ATLAS end-cap muon chambers (MEO and MEM) of 56 t being assembled in the cavern. Loads to be considered include the muon chambers, self weight of the structure and personnel mass. Different load combinations are calculated, not following specific regulations, neither accidental loading incidents. Boundary conditions include connections to cavern walls and floor and to the concrete shielding around the LHC beam exits into the cavern. Connections to the floor are always assumed fixed (all degrees of freedom blocked). In the other cases variations were calculated to include knee-joints instead of clamping, freeing the rotations.

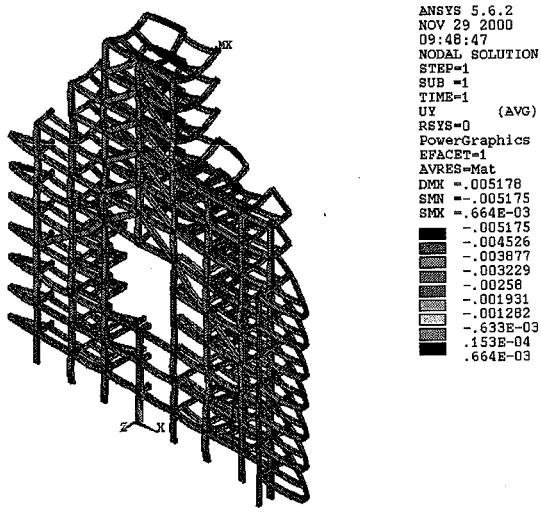


Figure 2.8: *FEM results for HO structure: Vertical displacements, [20].*

The results of the calculations show maximum displacements of $6-7\,\text{mm}$ in vertical direction for long horizontal beams. Differences between load situations considering the weight of the muon chambers or not have been found with a maximum of $2-3\,\text{mm}$ (not adding up to the maximum of before), thus the validation to carry the muon chambers was successful. For an example of the visualized results see Figure 2.8.

Conclusion for access structure deformation models The mechanical studies of the individual HS and HO structures did not consider any cavern deformation. The individual calculations consider certain load combinations which verify the structure considering maximum loads acting on it. The results show the maximum effect on the structures caused by these loads without including time as a factor in the model. For the application of these results to the expected deformations of the metrological cavern network points, it has to be considered that these points are all mounted on the vertical posts of the structures. For these posts only very small displacements ($< 1\,\text{mm}$) were found in the calculations. The results of calculations (although not being implemented in a dynamic model for deformation analysis) are used in a simulation of possible deformation scenarios of the ATLAS cavern network, see Chapter 6.3.

2.3 The ATLAS detector

2.3.1 Detector systems

The ATLAS detector is the largest experimental facility in the LHC project. The detector has a diameter of approximately 25 m and in closed configuration is approximately 40 m long. Its weight is about 7000 t. The main systems of the detector are Inner Tracking Detector (Pixel detector, SCT and TRT), Calorimeter (Electromagnetic and Hadron Calorimeter), Muon Spectrometer and the Magnet System (Central Solenoid and Toroid Magnet System), see Figure 1.2.

The main support systems are the feet based on the bedplates embedded in the cavern slab holding the Barrel toroid, and the main rails for the Barrel calorimeters housing the Solenoid and Inner detector systems. In the following some details about the detector subsystems are given. Details on the ATLAS detector summarized in this section and more can be found on the ATLAS project's webpage [3].

The Inner Detector measures the momentum of each charged particle. It combines high-resolution detectors at the inner radii with continuous tracking elements at the outer radii (see Figure 2.9), all contained in the **Central Solenoid** magnet, which provides a nominal field of 2 T. The highest resolution is achieved around the vertex region using semiconductor **Pixel detectors** followed by a silicon microstrip detector (**SCT** – Semiconductor Tracker). At larger radii tracking points are provided by the straw tube tracker (**TRT** – Transition Radiation Tracker). The outer radius of the Inner Detector is 1.12 m and the total length 7 m.

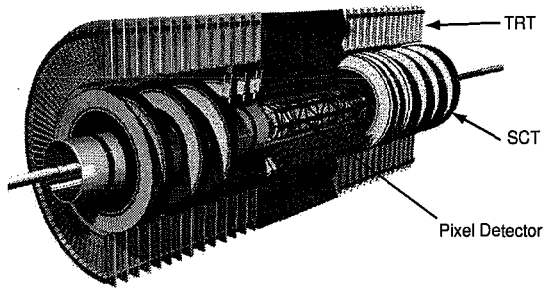


Figure 2.9: *The Inner Detector: Pixel Detector, SCT and TRT, ATLAS Experiment image,* ©CERN, [3].

The Calorimeter measures energies of charged and neutral particles. It consists of metal plates (absorbers) and sensing elements. Interactions in the absorbers transform the incident energy into a 'shower' of particles that are detected by the sensing elements. In the inner sections of the calorimeter, the sensing element is liquid argon (**Liquid Argon Calorimeter**). The showers in the argon liberate electrons that are collected and recorded. To obtain argon in liquid state, the gas has to be cooled to 88 K in a cryostat. In the outer sections, the sensors are tiles of scintillating plastic (**Tile Calorimeter**), see Figure 2.10.

The Muon Spectrometer identifies and measures muons. Muons are particles just like electrons, but 200 times heavier. They are the only detectable particles that can traverse all calorimeter absorbers without being stopped. The muon spectrometer surrounds the calorimeter and

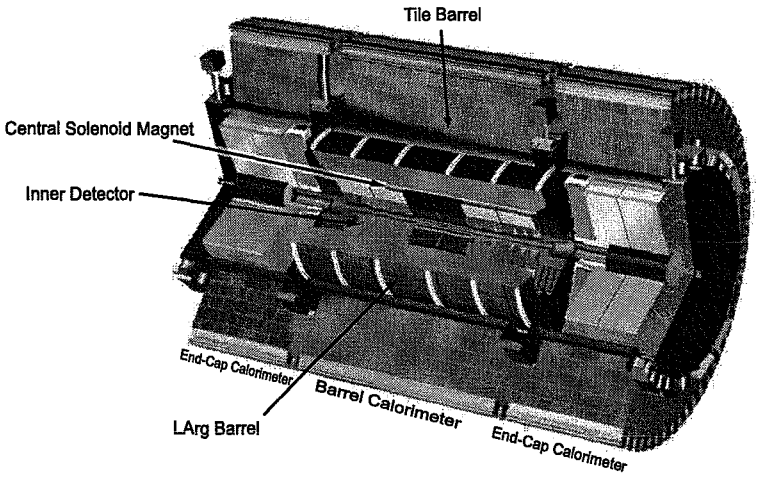


Figure 2.10: *The Calorimeter: Liquid Argon and Tile Calorimeters, Barrel and EndCap, ATLAS Experiment image,* ©CERN, [3].

measures muon trajectories to determine their momenta with high precision. It consists of thousands of charged particle sensors in individual **Muon Chambers** placed in the magnetic field produced by large superconducting toroidal coils (**Barrel Toroid** and **End-Cap Toroids**). The sensors are similar to the straws of the Inner Detector, but with larger tube diameters. There are more than 1000 muon chambers assembled inside the Barrel Toroid and around 1500 in the End-cap Toroids, see Figure 2.11.

The Magnet System bends charged particle trajectories to measure momentum. The **Central Solenoid** magnet has a length of 5.3 m with a diameter of 2.4 m. It is designed to provide a field of 2 T with a peak magnetic field of 2.6 T. The total weight is 5.7 t. The **Toroid Magnet** system consists of eight barrel coils housed in separate cryostats and two End-Cap cryostats housing eight coils each. The End-cap coil systems are rotated by 22.5° with respect to the Barrel Toroids in order to provide radial overlap and to optimise the bending power of the interface regions of both coil systems. The Barrel Toroid coils are assembled radially and symmetrically around the beam axis, see Figure 2.11. Each coil has an axial length of 25.3 m and extends radially from 9.4 m to

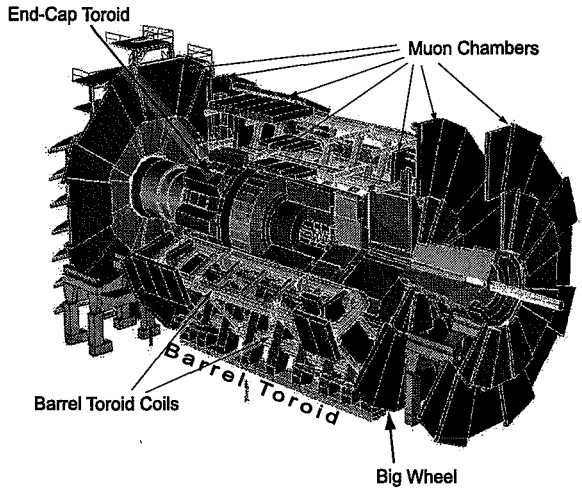


Figure 2.11: *Muon Spectrometer and Toroid Magnet System: Barrel Toroid and End-Cap Toroids, ATLAS Experiment image,* ©CERN, [3].

20.1 m. The total weight of all eight coils together is 830 t. The peak field provided by the Barrel Toroid coils is 3.9 T. The End-cap Toroid coils are also assembled radially and symmetrically around the beam axis. They are cold-linked and assembled as a single cold mass in one large cryostat. The cryostat rests on a rail system facilitating the movement and parking for access to the detector center. Each coil has an axial length of 5 m and extends radially from 1.65 m to 10.7 m. The total assembly weight is 240 t. The peak field provided by the end-cap toroids is 4.1 T.

2.3.2 Installation schedule

Here the planned installation schedule as outlined in [2] is briefly summarized. The real installation process of the ATLAS detector differs in some parts significantly from the planned schedule, both in sequence and time periods from this planned schedule. Nevertheless it gives a good overview of the major installation steps to be carried out. The progress of installation has to be considered in the design of the metrological network. The installed detector parts represent obstacles for the geodetic measurements. Information summarized in this section combined

with rough envelope specifications of the detector parts are included in a simulation of the ATLAS cavern network analyzed in Chapter 6.3.

Considerations in the installation schedule had to incorporate many different factors including necessary space, weight of detector parts to be lowered, sequence of the assembly, special transport constraints of individual systems etc. The heaviest object to be lowered is the End-Cap Toroid with 240 t. The biggest objects are the 8 individual coils of the Barrel Toroid each $25.3 \times 5.3 \times 1.1$ m large. The total installation period of the detector was planned to be about 30 months and the installation schedule was divided into 15 phases. The following summary is focused on major detector parts, often neglecting nevertheless important infrastructure installation. Survey interaction is demanded in all of the following steps but is not explicitly mentioned in the following description.

- Phase 1–3: After the cavern is delivered to CERN general facilities are installed including ventilation, power, lights, access structures and travelling cranes. With these cranes in place, enabling transport inside the cavern, the installation of access structures HO and HS starts and large cryogenic storage vessels are lowered into the cavern. Then the assembly of the ATLAS detector parts can really start with the installation of the support feet, see Figure 2.12 and the first Barrel Toroid coils.

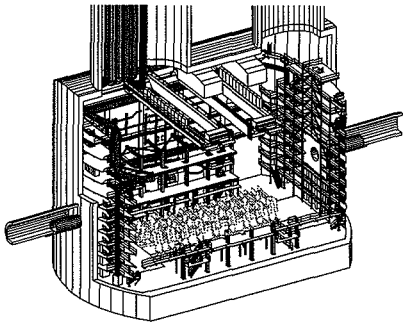


Figure 2.12: *Installation Phase 3: Start of the support feet and rail assembly, [2].*

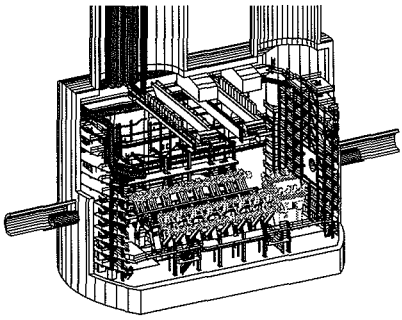


Figure 2.13: *Installation Phase 4: Barrel Toroid coils 3 and 4 in place. Start of Barrel Tile Calorimeter assembly on side C, [2].*

- Phase 4–6: More Barrel Toroid coils are installed and in parallel installation of the Barrel Calorimeter starts on side C, see Figure 2.13. The Liquid Argon Calorimeter Barrel is lowered into the cavern and assembled in connection with the Barrel Tile Calorimeter. Then the first Solenoid connections are installed and the Barrel Calorimeter is moved on rails to its final position inside the Barrel Toroid. The assembly of the End-Cap Calorimeter C can thus start on side C.

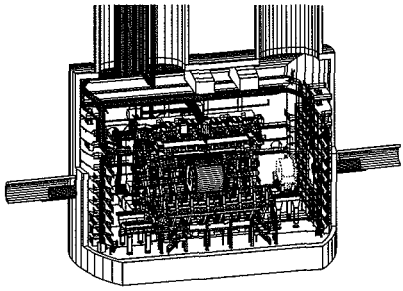


Figure 2.14: *Installation Phase 7: Liquid Argon End-Cap cryostat lowered onto supports Tile C, [2].*

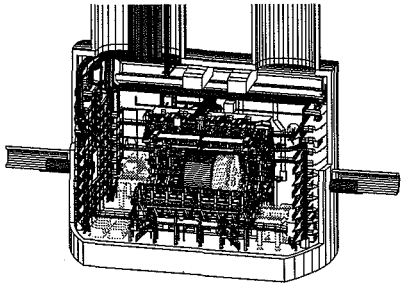


Figure 2.15: *Installation Phase 8: End-Cap Calorimeter and Shield Disc C. Start End-Cap Calorimeter A, [2].*

- Phase 7–9: The Liquid Argon End-Cap Calorimeter is installed on its supports in connection with the End-Cap Tile Calorimeter, see Figure 2.14. The End-Cap Calorimeter C assembly is completed and it is connected to the Barrel Calorimeter. The Shield Disc is foreseen to be lowered on side C, see Figure 2.15 and the End-Cap Calorimeter on side A is assembled as before on side C.

- Phase 10–12: The Shield Disc on side A is lowered. All services to the Barrel Calorimeter and the Solenoid are connected. After first tests of the Solenoid the installation of the Inner Detector can start. For this the End-Cap Calorimeters are moved out of the Barrel Toroid (open position). The barrel part of the Inner Detector is positioned as well as the first section of the beam pipe. Installation of Muon Chambers into the Barrel Toroid begins. The chambers are slid on previously installed rails inside the Barrel Toroid, see Figure 2.16.

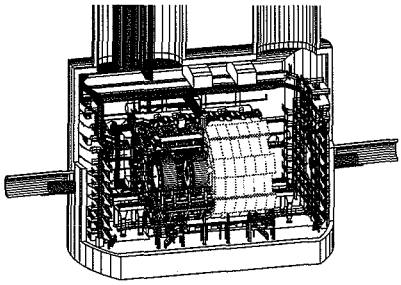


Figure 2.16: *Installation Phase 12: Start of muon chamber installation - barrel side C, [2].*

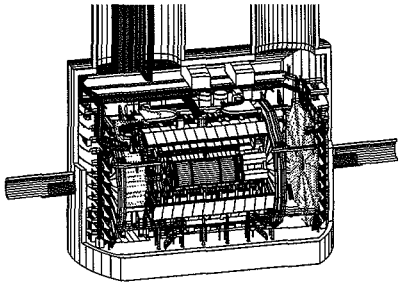


Figure 2.17: *Installation Phase 14: End-Cap Toroid A lowered onto support truck. Muon chambers side C, [2].*

- Phase 13–15: The Muon chambers on the Big Wheel on side C and A are installed. The End-Cap Toroids are lowered into the cavern. As they are very big, they are the last systems to be installed. They restrict access, particularly if in closed positions as this requires closing up all other sub-detectors. Their cool-down-time of 40 days is another restricting factor. During this period the installation of Muon Chambers on the wall structures can take place, see Figure 2.17. Finally ATLAS installation is completed and shielding blocks are lowered on both side A and side C. ATLAS is ready for taking collision data.

2.3.3 Detector access scenarios

When LHC is running, the experimental area (UX15 cavern) is not accessible due to radiation, [2].

Ultra-short shutdowns are planned to happen on a quite regular basis, allowing access to the experimental zone for a duration of approximately one hour. The detector is not open during this period thus no detector system is displaced. The magnetic fields stay on. Only equipment on the periphery of the detector is accessible in this type of shutdown. Access is only given to a very limited group of people. No survey interaction is foreseen for this kind of shutdown but if it becomes necessary a very careful and detailed planning has to be carried out beforehand.

In the *short-opening scenario* LHC and the detector systems are shut down and the experimental zone can be accessed under certain restricted conditions. The duration of these shutdowns is planned to be in the order of seven to ten days. The purpose of these shutdowns is to give access to some detector systems (muon chambers on the Big Wheels, Barrel muon chambers, space between End-Cap Calorimeters and End-Cap Toroids, outer faces of Inner Detector), thus some will be displaced, but major beam pipes stay in position and many service systems will stay connected. Survey interaction is only foreseen in very extraordinary cases and again needs to be planned in very high detail in advance.

Regular maintenance will be carried out during the shutdown periods of the LHC machine, scheduled every year between November and April for a period of approximately 5 months. During this period the ATLAS detector is in open configuration called *long-opening scenario.* This requires a large movement of the End-Cap Toroids. Some systems might even be craned to the surface (Forward Shielding, Shield Discs) to free space in the cavern. Nearly all moving system services will be disconnected except services on the End-Cap Toroid and on the Liquid Argon Calorimeter that have to be kept connected continuously to keep the system at cryogenic temperature. Survey measurements will take place to control structural deformation of the cavern especially the cavern floor, i.e. bedplates and to control and most importantly re-establish positions of detector systems before the next run.

2.3.4 Positioning requirements

The demands on surveys tasks in the context of the assembly and installation of the ATLAS detector are very diverse. To summarize, the steps of an individual detector part to be installed in the experiment are like follows:

- **Verification of manufacturing and pre-assembly precision:** This refers to both individual detector modules or lager structures that are assembled from individual detector modules. The precision is given by engineering specifications or manufacturing tolerances. It is in general verified outside the cavern before installation. Assembly precisions are very diverse and range from 30 μm

for some parts in the Inner Detector to 5 mm for the Barrel Toroid coils, corresponding survey measurement precision demands vary between 50 μm – 1 mm.

Different types of information requested in this step include dimensions of objects, planimetry and relative alignment of individual modules. The geometry of the detector parts are generally represented by at least a minimum set of selected points (fiducial marks) on the outer surface of the object which are visible and accessible in the installation procedure. These fiducial marks are in general equipped with a standard reference hole to fit survey targets commonly used by the CERN experimental survey group.

- **Accuracy of installation with respect to the nominal beam line:** The installation of detector parts is guided by survey measurements using the fiducial marks or other features of the object.

  The accuracy of the positioning of the individual detector parts is effected not only by survey measurement precision and accuracy in the metrological network but also by manufacturing precision and mechanical assembly constraints. Overall it can be said that the necessary accuracy for survey tasks is between 50 μm – 0.1 mm in the area of the Inner detector and between 0.5 – 2 mm on the outer.

As one example for better illustration of the diverse requirements may serve the Inner Detector, [18]:

The Inner Detector sits on rails attached to the inner diameter of the Liquid Argon Cryostat of the Inner Warm Vessel. These rails have to be aligned to be planar to ± 0.1 mm and parallel to ± 0.2 mm. The absolute position of the Inner Detector with respect to the nominal beam line depends thus on the accuracy in the metrological network and also on the positioning accuracy of the Liquid Argon Cryostat warm vessel. It is required to be ± 1 mm. The error budget for the survey accuracy is 0.1 – 0.3 mm.

The Inner Detector and the Liquid Argon Calorimeter are located in the very center of the ATLAS detector and are installed when the Barrel Toroid is already in place, thus access and network configuration are very difficult. A good reliability in the network is essential to give accurate positioning results. Special measures are taken (installation

of temporary additional network point monuments, etc.) in order to give a good connection to the network. The relative positioning of the individual detector systems in the Inner Detector are in the range of $10-1000\,\mu$m, the most central pixel detector being the most demanding, and less for the TRT.

When methods available to the survey group are not sufficient or the required access can not be given, relative positioning has to be achieved by other methods including services provided by a metrology lab or specially designed alignment systems.

Positioning requirement details for any system can be found in the TDR (Technical Design Report) of each system in the Electronic Document Management System (EDMS) at CERN.

2.4 The metrological ATLAS cavern network

A metrological cavern network was originally designed in a preliminary study considering the following constraints, [13]:

- Space with respect to construction specifications and needs to perform geodetic measurements ensuring personal safety,
- requirements for spatial uncertainty (< 0.2 mm (1σ) for each point),
- local redundancy of observations ($> 60\%$),
- distance measurements are preferred over angle observations in order to minimize measurement uncertainty and effort.

The resulting *nominal network* design consists of 116 point monuments in the cavern, installed along the cavern walls, embedded in the floor and mounted to the HS and HO structures. This design study has been used as a guideline for the installation of the real network point monuments in the cavern but constraints that arise in the real situation cause the real network to differ in some places substantially from the design layout.

Two figures illustrate some network point monument locations in this design network: In Figure 2.18 the monuments distributed on the HO structure on Side C can be seen. These monuments are plug-in

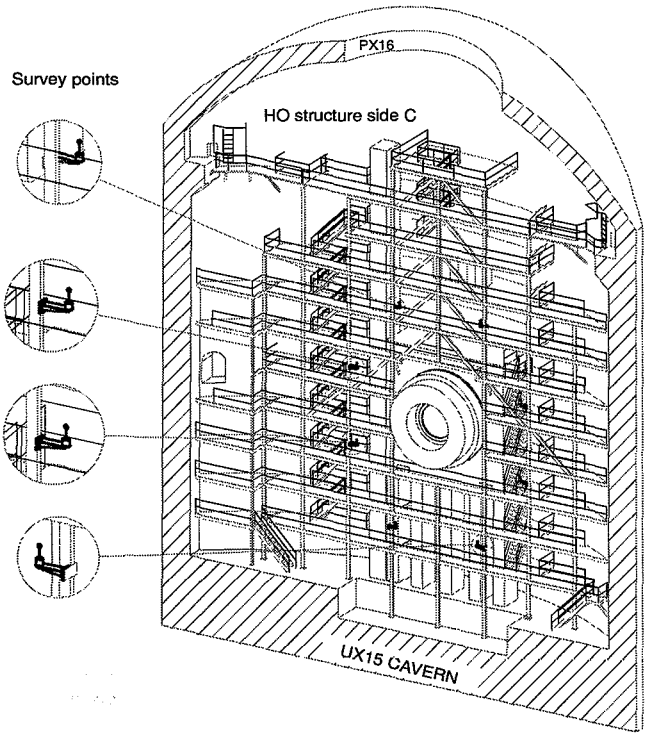


Figure 2.18: *Network point monuments on HO structure Side C. Plug-in brackets mounted on vertical posts of structure in different orientations,* ©CERN–TS, Ref. ATFIM___0002.

brackets, see more details about CERN point monumentation in Chapter 3.2. In Figure 2.19 network point monuments along the US cavern wall are presented. The distribution considers the available space not interfering with infrastructure or other installations (not plotted here) and the necessary space to actually carry out theodolite observations from such a point, indicated by the spatial envelopes around the monuments. Additionally, access to the point monuments from gangways and access structures should be possible and the demands on the network summarized above should be fulfilled.

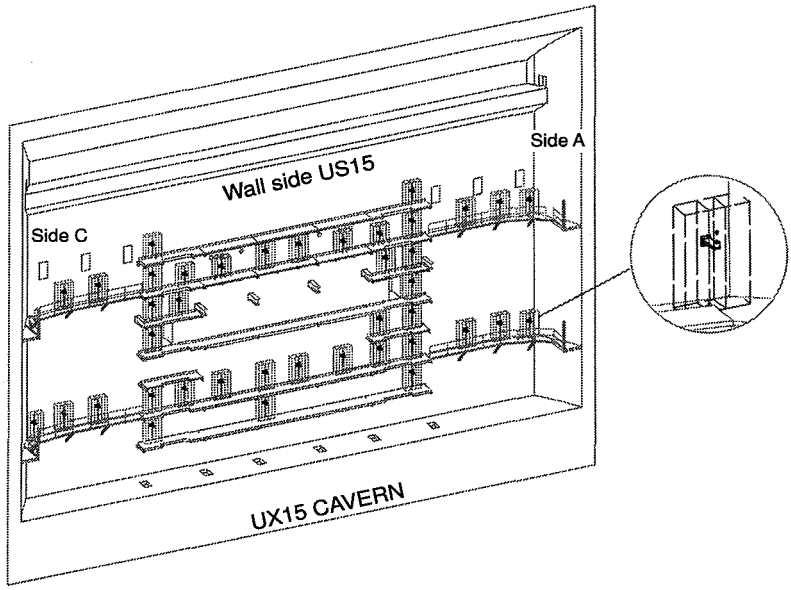


Figure 2.19: *Network points on cavern wall Side US. Foldable brackets attached to walls and the envelope of the necessary space around to make observations form these points,* ©CERN–TS, Ref. ATFIM___0008.

A simulation of measurements between the cavern network points inside the cavern was used to verify the network configuration. However, this simulation did not consider some important factors:

- Datum definition: The ATLAS cavern network refers to LHC geometry (represented by reference points in the LHC tunnel). Points in the LHC tunnel are not directly visible from the cavern after a certain stage of installation thus the only connection is achieved passing through the survey galleries. The link through these small galleries is geometrically very weak and thus it will be realized by special metrological systems: A hydrostatic levelling system (HLS), a wire positioning system (WPS) and additionally a precise distance measurement system consisting of calibrated invar bars and capacitive sensors. See Section 3.3 for more information on these special measurement systems.

The datum definition is transferred by these systems into the cavern. As these systems are foreseen mainly to support LHC alignment (control of Inner Triplet quadrupoles) they are not available from the beginning of detector installation, see layout in Figure 2.20. Thus before these systems are available the connection to the datum definition is achieved directly via the tunnel exits (as longs as these are not blocked) or through the critical survey gallery link.

- The progress of installation affects the network configuration severely, as many sightings that are possible at the beginning become obstructed as more objects are installed.

A simulation of survey data considering the installation process and also possible deformation scenarios is used in Section 6.3 to analyze the nominal network layout and to evaluate the performance of the data processing algorithm presented in Chapter 5.

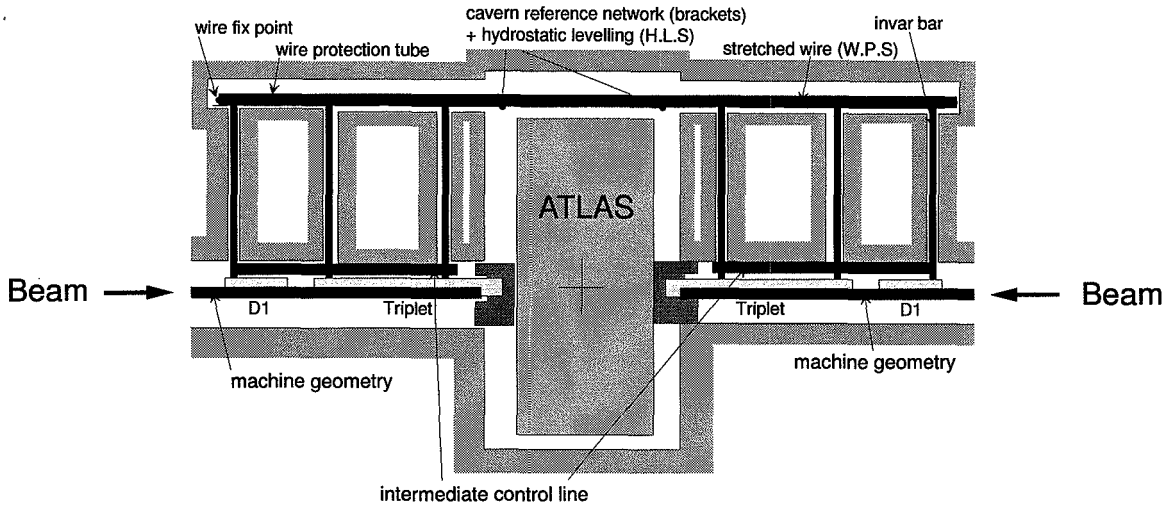


Figure 2.20: *Schematic top view for the installation of special measurement systems HLS, WPS and invar bars (IRS) in the LHC tunnel and ATLAS experimental zone, [2].*

3

Metrology at CERN applied to detector positioning and alignment

Ever since CERN was founded more than 50 years ago, accelerator and experimental installations needed to be positioned and aligned. The challenges in these applications have always demanded for new and creative surveying solutions. Alignment techniques and high-precision measurement systems developed and employed at CERN and other accelerator laboratories are different to other applications in engineering surveying.

In this chapter an overview is given of techniques developed and used today at CERN in the context of metrology for high energy physics experiments as well as some details about geodetic reference systems and corrections. Special focus is given to a selected group of high precision measurement systems employed in the ATLAS cavern network. Although the development of measurement techniques or systems is not

part of the research work presented in this thesis the purpose of this overview is to illustrate the survey work and special developments carried out in the CERN survey group over recent years. Most of the methods presented in the following are employed in the real ATLAS cavern network data set analyzed in Chapter 6.4.

For more general details on accelerator alignment and related surveying tasks reference is made to Chapter 4.1 in [33].

3.1 CERN reference systems

The CERN reference system is a local geodetic reference system represented by the CERN Coordinate System (CCS), [23]. It is a three dimensional Cartesian coordinate system. Its principal point is a survey point in the center of the old PS accelerator, with the Z-axis being the vertical at this point, P_0. The X and Y axes are defined locally forming a right-handed Cartesian coordinate system, see Figure 3.1.

Part of the geodetic reference system are definitions for a horizontal and a vertical geodetic datum. The CERN horizontal geodetic datum is defined by a geodetic reference ellipsoid, in the form of the GRS80 reference ellipsoid. Its relative position and orientation are defined by the geodetic coordinates of P_0, setting the deflection of the vertical at this point to zero and fixing the geodetic azimuth of the CCS Y-axis in the direction of another real PS survey point.

Additionally to the horizontal datum four different vertical datums are used at CERN. The complexity of the related models depends on the size and accuracy requirements of the accelerator project involved, [22]: Whilst for the PS accelerator with a diameter of 200 m a planar reference system sufficed, it was necessary for the SPS accelerator (diameter of 2 km) to establish a spherical reference surface.

For the LEP accelerator (diameter almost 9km) a local geoid model was derived. In a collaboration with Swiss academic and federal institutes a high resolution gravity field model was simulated for the CERN site and its results compared to astro-geodetic measurements. This geoid model is represented by a hyperbolic paraboloid and describes the equipotential surface of the gravity field passing through point p_0 which is the foot of the CERN ellipsoidal normal passing through the

principal point P_0, see Figure 3.1. This model was also to be used for the LHC which is installed in the very same tunnel as LEP before.

But for the CNGS Project – CERN Neutrinos to Gran Sasso (730km from CERN) – it was decided to update the geoid model over the CERN site, [24]. In collaboration with the Swiss Federal Office of Topography, values for the deflection of the vertical were extracted from the model used there (an evolution of the models used at the time of LEP) for a grid of points over the CERN site. As no analytical surface could provide a good fit to the values obtained, interpolation is carried out using splines. The new geoid model is referred to as RS2K and is applied to calculations for LHC today as well.

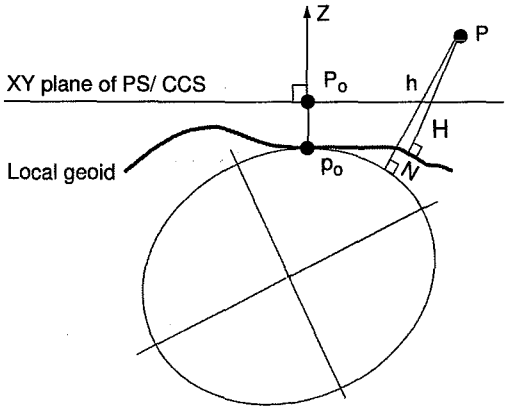


Figure 3.1: *CERN Coordinate System defined at the initial point P_0. Relationship to reference ellipsoid and local geoid model.*

With these local geoid models it is possible to derive deflections of the vertical and geoidal undulations for any point at CERN with respect to the geodetic reference ellipsoid. Furthermore it becomes possible to combine horizontal and vertical positions to get 3D positions and to treat measurements together and not, as common in traditional methods, to separate horizontal and vertical networks in calculations. Similarly the relationship to any local geodetic reference system (local survey reference system) can be defined. Simplified transformation routines for all LHC interaction points can be found in [30].

Thus, with the information of a local geoid model, survey data can be adjusted and analysed in the CERN Coordinate system (CCS). To achieve this the observation equations are determined in the local geodetic system and the resulting matrices are transformed to the CCS for adjustment.

Survey measurements inside a particular experimental zone are calculated in the respective local, horizontal experiment coordinate system, like the ATLAS survey coordinate system described in Section 2.1. The transformation parameters between the CCS and the ATLAS survey coordinate system defined for the LHC Point 1 can be found in [30].

Survey data inside such a local reference system are analyzed without corrections for effects of the earth's curvature. Nevertheless network measurements linking the local network to the LHC geometry and thus the CERN reference system are calculated in general including the respective corrections. If the earth's curvature is not taken into account in the local system the resulting error estimate for a distance of 50 m is approximately 0.2 mm. This number (x) is obtained by a spherical approximation of the geoid surface with a mean earth's radius of $R = 6371\,\text{km}$ and a length $L = 50\,\text{m}$, see Figure 3.2.

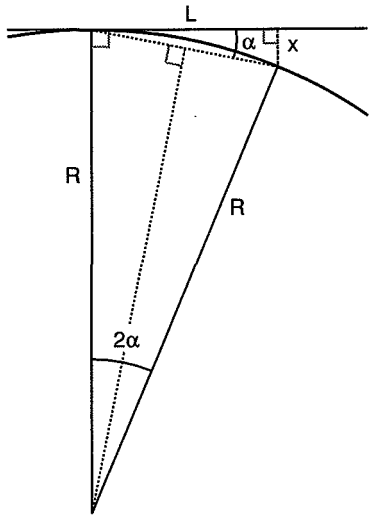


Figure 3.2: *Estimation of error for neglecting earth's curvature.*

Note: The scope of this thesis is to present an algorithm solving deformation problems occurring in metrological networks and not to provide a complete survey data adjustment package. The survey data adjustment module providing data for the adaptive Kalman filter algorithm presented in this thesis does not apply any corrections for the shape of the earth. This algorithm is supposed to be implemented as part of an survey data analysis program that provides the above discussed corrections if necessary, like the least-squares adjustment programme LGC [31] developed and used by CERN's survey group. Thus the results for real network survey data presented later in this thesis are adjusted in a horizontal coordinate system, not considering any corrections for the deflection of the vertical. More details are given in Chapter 6.

3.2 Network point monumentation

In high precision metrology special care needs to be taken in the point monumentation in order to minimize centering and repeatability errors. To further enhance repeatability, point coordinates generally refer to actually measured points.

A reference surface that can be very precisely manufactured and its position reproduced, is a sphere. CERN reference points are generally realized by reference sockets, see Figure 3.3.

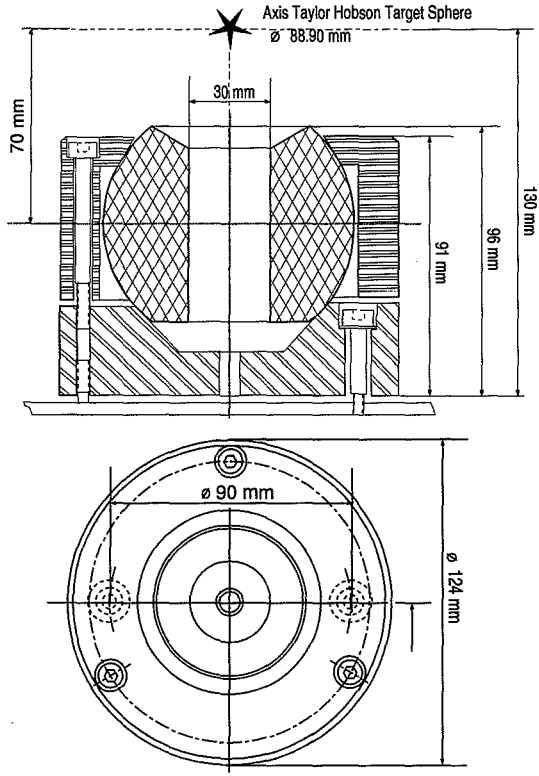


Figure 3.3: *CERN survey reference socket. Vertical section (top) and top view (bottom).*

These reference sockets are based on a simple principle: A sphere touches a coaxial conical surface along a circle. The very basic reference point is thus simply a base incorporating the conical surface on which a precisely manufactured Taylor-Hobson sphere is installed, see Figure 3.4 (top left) and (top center). The sphere touches the conus along a well determined circle. If the sphere is rotated, it still touches the conus along that very same circle. The repeatability of this position is very precise ($< 10\mu$m) but it can only be used for targets.

To allow also the installation of an instrument referring to this point the principle of sphere and coaxial conical surface is employed another time. In the standard reference socket a sphere which serves as extension is positioned on the conical base surface, their axes coinciding. The sphere itself has a hole with a diameter of 30 mm. On the top this cylinder evolves into another coaxial conus, see Figure 3.3. The sphere can be fixed in its position by a cap touching the sphere along a circle. This cap is screwed to the base by three screws. The reference socket can be levelled by inserting a tool into the cylindrical hole which holds a precise spirit level. The sphere inside the socket is then rotated until its axis is aligned along the local vertical.

The reference socket is then used to install either a Taylor-Hobson target, a similarly shaped prism or to station an instrument, see Figure 3.4 (bottom). A special base can be inserted into the socket. This base is either equipped with a standard Wild screw that is fixed to the Leica standard instrument base. Alternatively a Wild *GDF* 3-pin theodolite base (as used for the T2) accepting a theodolite is directly fixed to the insert part, see Figure 3.5, which enables a very precise centering of the instrument of $< 20\mu$m.

In a working environment like an experimental zone or also the accelerator tunnel it is important that survey reference point monuments can be protected from any external forces. This is achieved in some cases by removing part of the monument structure or by embedding monuments in the floor and applying additional protection measures.

In the following standard reference point monumentations are described and examples for application in the ATLAS cavern network shown.

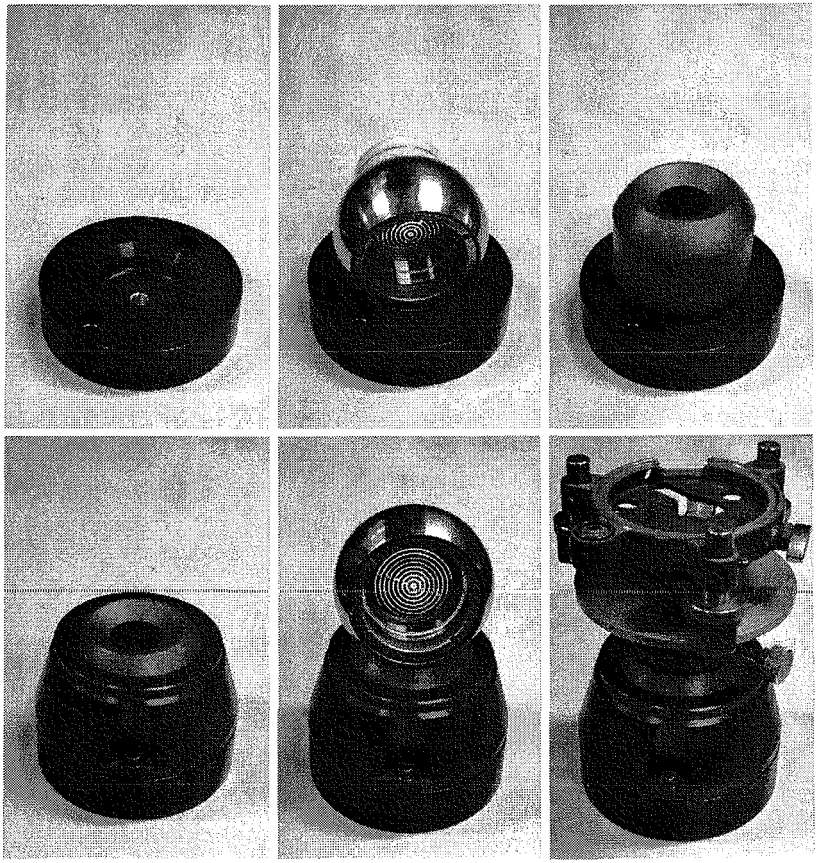


Figure 3.4: *CERN survey reference socket. Base only (top left), base with Taylor-Hobson target (top center), base with sphere (top right). Completely assembled socket with cap (bottom left), socket with Taylor-Hobson Target (bottom center) and socket with theodolite base (bottom right).*

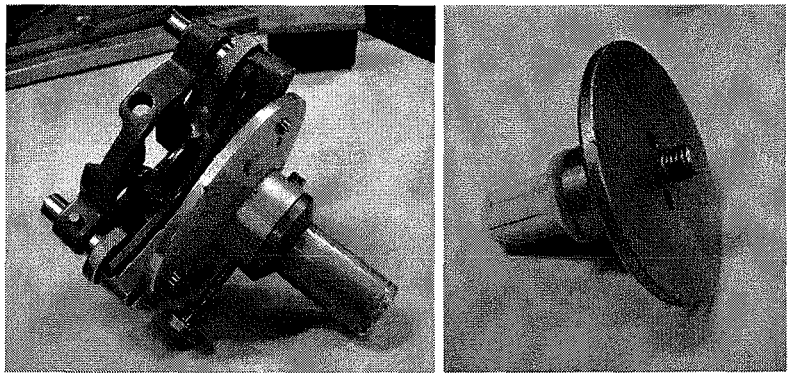


Figure 3.5: *Instrument base fitting into standard socket. Wild GDF base (left) and Wild standard screw (right).*

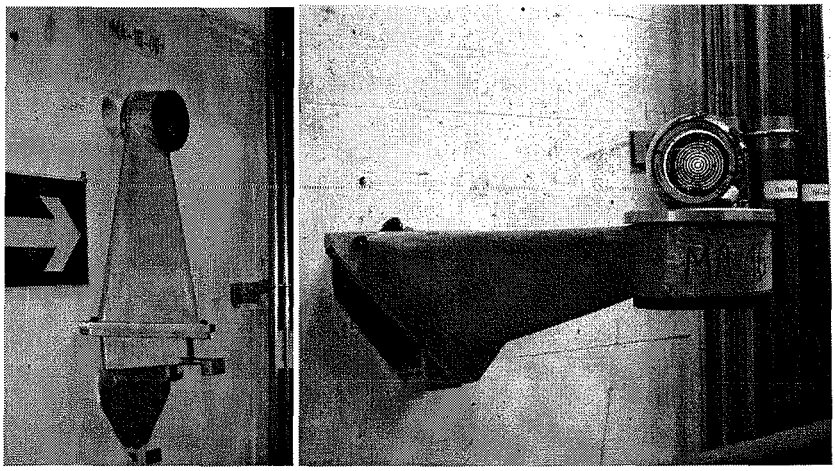


Figure 3.6: *Foldable reference point bracket. Bracket folded up (left) and in position for use, here equipped with close-distance Taylor-Hobson target sphere (right).*

Foldable brackets These brackets are used for wall points along the cavern walls, see Figure 3.6 and Figure 3.7. They had already been used in LEP experiments and proved a repeatability of $< 50\mu$m. These brackets need very little space when not in use and are quickly installed.

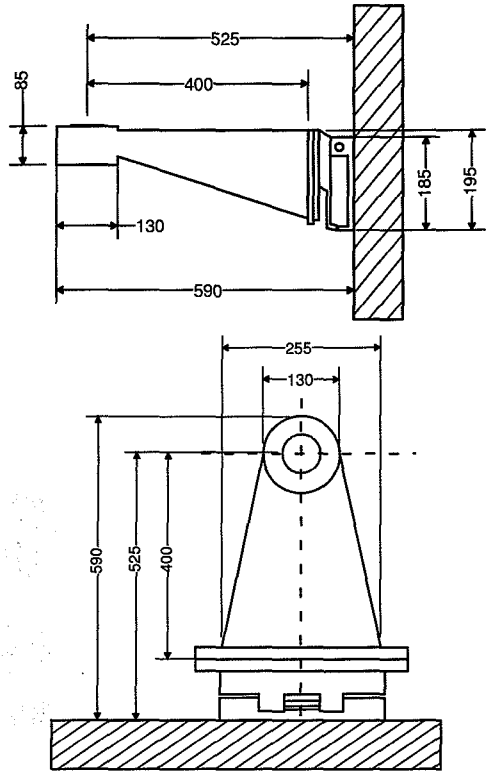


Figure 3.7: *Foldable reference point bracket. Side view (top) and top view (bottom). Dimensions are given in mm.*

The work position of the foldable bracket is achieved by unscrewing a security pin and folding the bracket down. The reference socket for this type of bracket is similar to the standard socket by means of the sphere with coaxial cylindrical hole and conical top finish. As before it can be levelled by rotating the sphere in its three-screw-fixation.

Plug-in brackets These brackets are used in the ATLAS network mounted on the access structures HS and HO or other structures. They are completely removed when not used and stored in safe areas close-by, except a baseplate which remains in place. The brackets referring

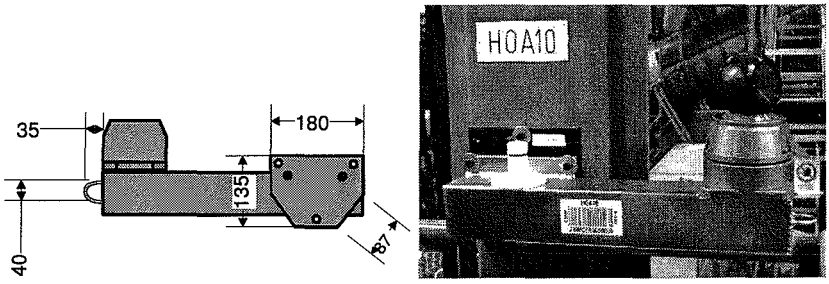


Figure 3.8: *Plug-in reference point bracket – type right. Dimensions in mm (left) and picture of bracket mounted on HO access structure in the ATLAS cavern (right) equipped with prism in sphere.*

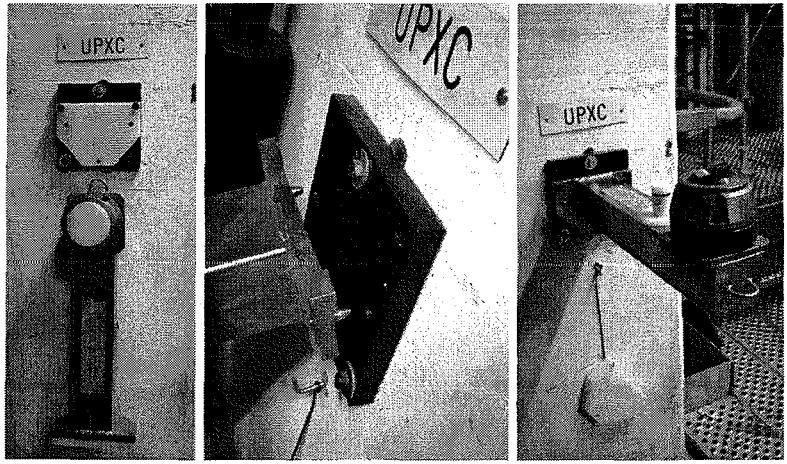


Figure 3.9: *Plug-in reference bracket – type forward. Left: Bracket stored, baseplate and plug-in surface protected by plastic covers. Center: Bracket pins are inserted into baseplate. Right: Bracket in work position.*

to different points should not be interchanged, to ensure repeatability. These brackets exist in different types depending on the direction they are fixed to the metal structures.

The base plate is screwed to the metal structure or a wall. Screws are also used to fix the bracket to the base plate. When installing, first two pins on the bracket are inserted into each one hole and one slot (allowing horizontal movement only) in the baseplate until the bracket touches the baseplate on three defined points. This is done so that these pins take the load of the bracket weight instead of the later inserted screws and to avoid any strain on the bracket itself. See for illustration Figure 3.8 and Figure 3.9. These brackets have been found to reproduce a reference point's position within $< 50\mu$m even under a load of 15 kg. They are directly equipped with a standard reference socket.

LHC tunnel reference points These monuments stem from the period of the LEP project. They are embedded in the floor in special pots below floor level, covered by a protection plate, see Figure 3.10. A metallic insert finishes with a conical surface ending in a cylindrical hole below. This insert accepts an similarly shaped extension of 70 mm, with a cylindrical base, leading into a conical shape, housing another cylindrical hole and finishing in a conical surface. On this surface then a Taylor-Hobson sphere target can be positioned, its center representing the reference point to which coordinates refer.

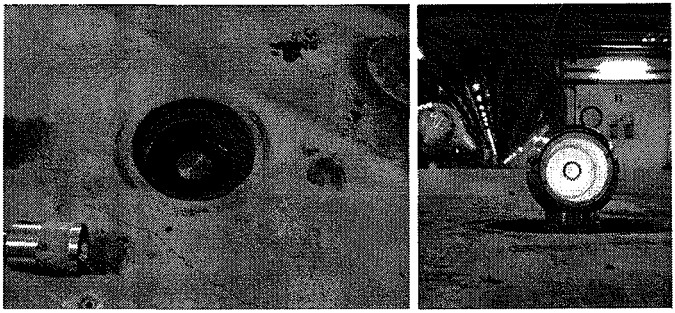


Figure 3.10: *LHC tunnel reference monument embedded in tunnel floor. Protected by plate (level with floor) and plug. 70 mm extension to be inserted into this monument (left). Extension added and Taylor-Hobson sphere for long-distance observation positioned (right).*

For setting up an instrument station above this type of point special portable survey columns are used, see Figure 3.11. Such a survey column has a spherical pivot insert on its base. Its three feet are used to align

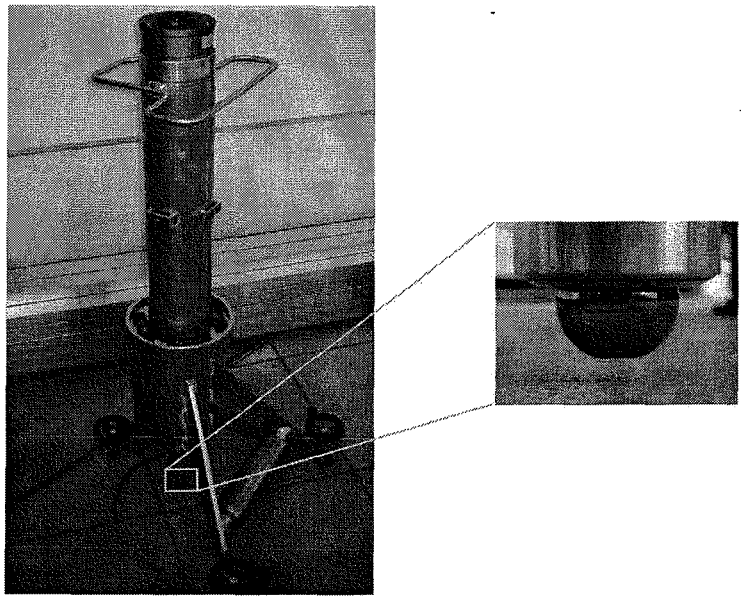


Figure 3.11: *Survey column for LHC tunnel reference points (left). Pivot sphere on the bottom (right).*

the column with respect to the local vertical. Precise manufacturing of all mechanical parts again ensures that all of these elements are coaxial, resulting in a very precise centering and repeatability of the reference point's position of $< 10\mu$m.

Cavern floor points These monuments are embedded in the ATLAS cavern floor. The small reference supports are generally screwed to the base slab concrete approximately 3 cm beneath the floor level and are protected with a lid, level with the floor, see Figure 3.12. In some special locations close to the cavern corners, monument bases have been installed before the concreting of the final floor level and small reference sockets fixed

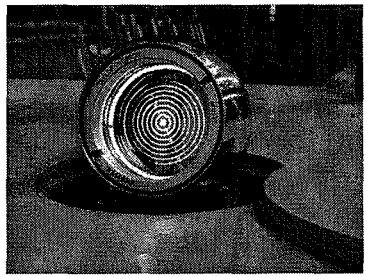


Figure 3.12: *Reference point monument embedded in cavern floor. Protection plate level with floor.*

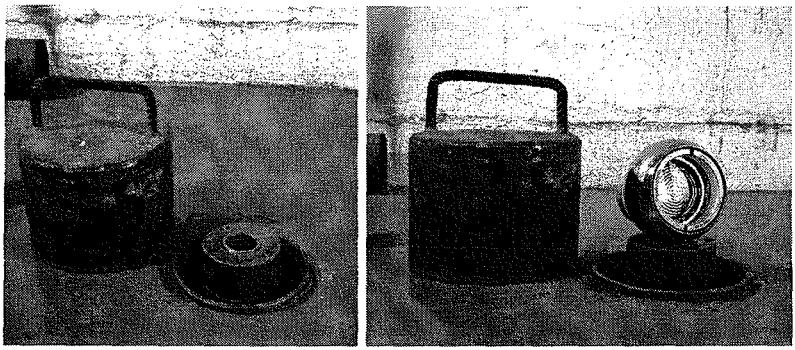


Figure 3.13: *Reference point monument embedded in cavern floor. Monument with cup protection (left) and equipped with Taylor-Hobson sphere (right).*

to them. In these cases sockets are not level with the floor and have to be protected by robust caps, see Figure 3.13. Repeatability for this type of point realizations in combination with a Taylor-Hobson sphere is $< 10\mu$m.

Deep reference points in LHC tunnel These type of points are actively linked to the bedrock by a rod in a depth of $20 - 30$ m. These monuments serve as height reference and are regularly monitored. The link to levelling measurements is achieved by a rod with a conically shaped surface finish end piece, accepting levelling staff feet, see Figure 3.14. A cover similar to LHC tunnel reference points is used to protect these monuments. They are located in the LHC tunnel, one or two in the vicinity of each LHC interaction point.

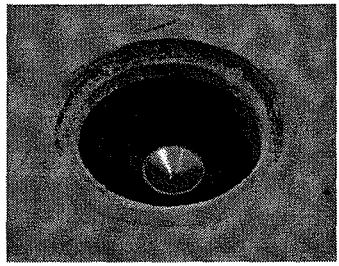
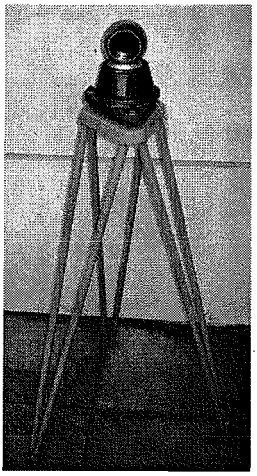


Figure 3.14: *Deep reference embedded in bedrock below LHC tunnel.*

In the case of the ATLAS experiment two points are available, one on each side of the cavern.



Points on plug–in tripods These monuments are installed on steel tripods, fixed with three screws to the floor, see Figure 3.15. Special brass plugs in the floor secure the screws. A standard reference socket is fixed to the tripod and thus an instrument can be stationed. These tripods are easily installed and also removed if necessary, the repeatability is within $< 100\mu$m.

Figure 3.15: *Tripod with Taylor-Hobson target.*

3.3 Measurement methods

In order to accurately position detector parts and other installations, different measurements methods are employed. In this section a short summary is given of the various types of measurements used in the ATLAS cavern network. Special focus is given to non-standard systems which have been specially developed or adapted for this application.

The standard instrument to carry out classical tachymetry is the Wild TC2002 instrument, in combination with Taylor-Hobson sphere targets. Angle and distance measurements are carried out separately, exchanging the target with a corner cube prism. Measurements are controlled on-line using a special software module running on a portable computer. The data is immediately corrected for atmosphere and prism constants and stored.

Leica Laser Tracking Devices (LTD500) are also available in CERN's survey group to carry out precise point positioning (< 0.1 mm). This tracker's portability is limited as are possible applications inside a cavern network. Instrument stations would only be possible on the cavern floor due to the instrument's weight and demands for a stable platform. Thus it is not employed in the cavern networks and is not discussed here in detail.

Spatial distance measurements

Electro-optical distance measurements are part of classical tachymetric measurements, carried out generally with Wild TC2002 instruments. A precision of $< 0.3\,\text{mm} + 1\,\text{ppm}$ for a distance measurement to a prism is achieved. Additionally two KERN Mekometers ME 5000 are available with precision of $< 0.2\,\text{mm} + 0.2\,\text{ppm}$ for distances $> 3\,\text{m}$ but are rarely used.

Atmospheric corrections are applied to the distance measurement directly taking into account measured temperature and pressure corresponding to formulas provided by the manufacturers. Generally only temperature measurements at the instrument station are taken, as the distances are short and the atmospheric environment is considered stable in this area. Instruments are regularly calibrated on the CERN interferometric calibration base in certain reflector combinations. The appropriate prism constants are then applied to every distance measurement.

Methods employing Distinvar (an automatic invar wire measurement system developed at CERN) have not been heavily used in recent years for measurements in the experimental caverns. These systems are far less flexible in their use than electro-optical distance measurement systems, as the invar wires are always of a fixed length and measurements are only possible for approximately horizontal lines. Nevertheless, these measurement methods are very precise ($< 0.1\,\text{mm}$) and are thus not easily replaced where their application is necessary.

Vertical distance measurements - geometrical levelling

Precision levelling is carried out generally with either a Wild N3 or NA2 level with micrometer plate achieving an accuracy of $< 0.2\,\text{mm}$ and $< 0.3\,\text{mm}$ for 1 km double path levelling, respectively. The instruments are regularly checked on a collimator in the laboratory. The NA2 minimum focus range is 1.6 m. If shorter sightings have to be done, a Wild N3 instrument has to be used. Its shortest focus range is 45 cm.

Invar-levelling staffs are equipped with special feet to reproduce the feet's position precisely, similar to the reference sockets described above and also to achieve the interface between levelling and 3D measurements without any loss of accuracy, see Figure 3.16. The foot is dimensioned

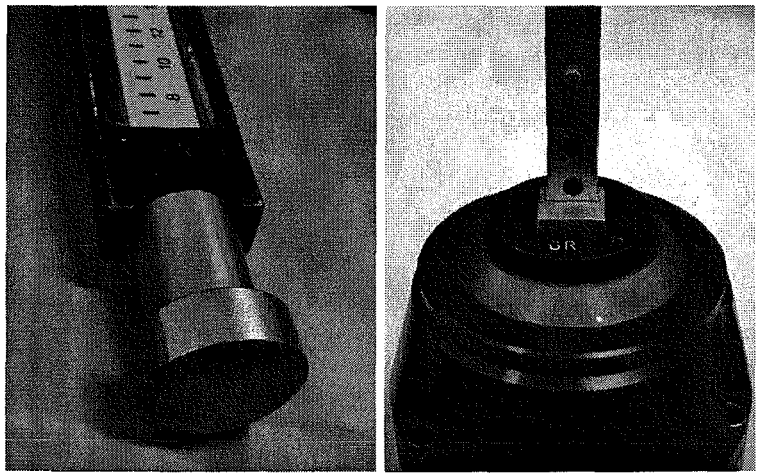


Figure 3.16: *Foot of standard levelling staff (left) and small ruler used for levelling inserted into standard reference socket(right).*

to set the zero level of the staff in exactly the same height as the center of the Taylor-Hobson sphere when inserted to the same reference monument.

The reason for using traditional levelling instruments and not instruments of the new, digital generation lies in the diversity of applications. For a digital level to measure precisely it is necessary that a major portion of the levelling staff is visible from the instrument. In an environment like experimental installations this can often not be achieved as objects may obstruct the sight to parts of the levelling staff. Also it is sometimes necessary to use very small staffs or even rulers with lengths of down to 20 cm.

Horizontal and vertical angle measurements

Angle observations are carried out with Wild TC2002 instruments on both faces for network measurements. Each single measurement is carried out twice and the average value stored, see Figure 3.17.

For angle measurements a Taylor-Hobson sphere with an illuminated target of concentrical rings is used. Target patterns are available in different types, to be better visible in different situations, especially for different distance ranges. The measurement precision of a single reading is $< 1.5^{cc}$ horizontally and vertically. The field procedure requires each set of angle measurements to be finished by a closure (remeasuring the first point again at the end) in order to check for any instabilities during the set of measurements. The instruments' collimation is regularly checked in the field.

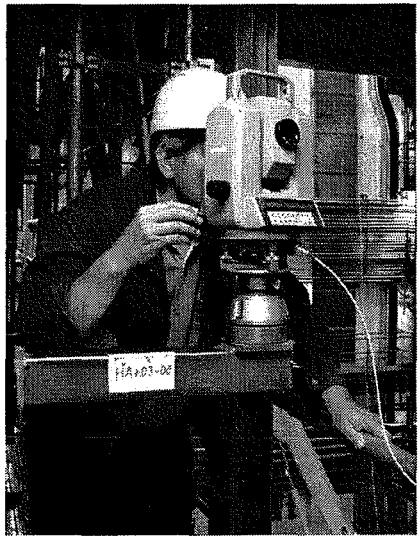


Figure 3.17: *Angle and distance observations carried out with a TC2002 in the ATLAS cavern.*

Wire Positioning System - WPS

The Wire Positioning System is one of the new special survey systems that have been applied at CERN in recent years in order to solve high precision survey tasks in the LHC and also for a new generation of linear colliders like CLIC (Compact Linear Collider) which might follow after LHC.

In the vicinity of the ATLAS cavern the systems WPS, HLS (Hydrostatic Levelling System) and IRS (Invar Radial System) are used to establish a connection between both sides of the LHC accelerator through the experimental zone. The alignment of these last sections with respect to each other is very critical and has to be monitored continuously. These systems are also used to transfer the LHC geometry realized by point monuments on the magnets into the cavern network and thus also the datum information. For the overall layout see again Figure 2.20.

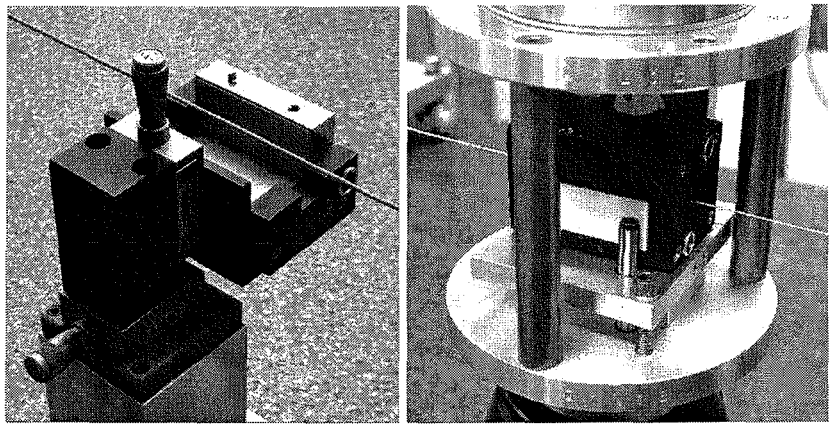


Figure 3.18: *Wire positioning system: Sensor and stretched wire, (left) in open position and (right) in closed position installed on reference support with interface to HLS system.*

WPS measures offsets of a stretched wire in two directions orthogonal to the wire by the principle of a capacitive sensor, [9]. The wire passes the WPS which has the form of rectangular parallelepiped, see Figure 3.18. The wire representing the spatial reference is actually a twisted carbon composite fibre wire. Together with the sensors made out of metal and ceramic a variable capacitor is formed. The sensors used at CERN are manufactured by *Fogale nanotech* and have a measurement range of $\pm 5\,\text{mm}$ and a precision of $< 5\mu\text{m}$. The sensors have two electrodes per axis which gives very precise results. If necessary the electronics connected to the sensors can be some tens of meters away which is necessary in high radiation areas. The wire is stretched with an appropriately dimensioned counterweight. Precise modelling of the catenary (considering the length of the wire, its weight, the counterweight and the height differences between wire endpoints and wire mid-point) is necessary in order to achieve the maximum precision.

The WPS system is installed on the LHC Inner Triplet quadrupoles on both sides, [10]. Its main objective is to monitor the horizontal radial alignment of these magnets with respect to each other. To achieve the link of both sides an Offset Reference Line passes through the survey galleries and the experimental cavern over a length of 126 m, see Figure 3.19. The connection between the Inner Triplet Lines and the

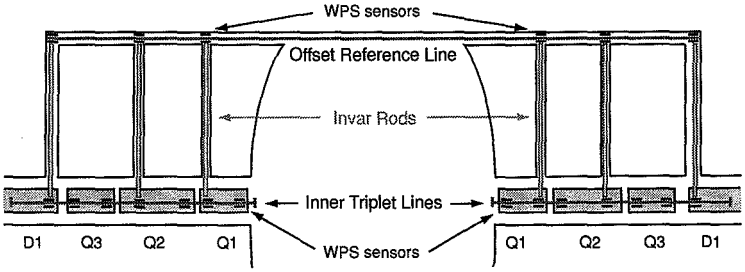


Figure 3.19: *Layout of WPS and IRS installation around ATLAS cavern.*

Offset Reference Line is achieved by a radial system of invar rods (IRS) composed of three rods on each side of the cavern.

Invar Radial System - IRS

The invar rods with a length of 16 m are installed radially between the survey galleries and the LHC tunnel in 12 m long boreholes of 40 cm diameter, [34]. They are calibrated beforehand at CERN's interferometric calibration base. Each end of the invar rods is equipped with sensor targets, see Figure 3.20. Capacitive sensors that are fixed to WPS sensors measure the distance to these targets continuously. By applying atmospheric corrections the offset of the WPS Inner Triplet Line and the Offset Reference Line in the horizontal plane is thus known with a precision of $< 100\mu$m.

Hydrostatic Levelling System - HLS

In vertical direction the link between both sides of the LHC Inner Triplets is realized by a closed hydrostatic loop with one step. The step is necessary as the inclination of the beam over the length of the hydraulic pipe exceeds the measurement range of the sensors. Thus two individual hydraulic pipes are used. At the step, two HLS sensors are linked mechanically, thus the offset is well determined. The HLS is installed passing the survey galleries and the experimental zone, see Figure 3.21. The combination of this HLS with the WPS and IRS determines the radial geometrical component of this link very well. The

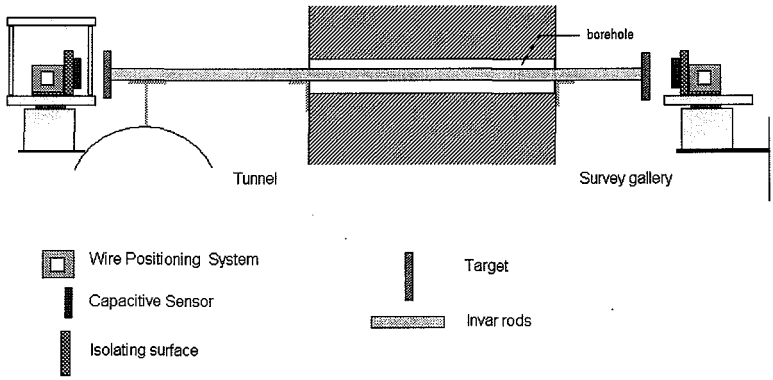


Figure 3.20: *Invar Radial System IRS linking stretched wire lines in the LHC tunnel and survey galleries, [10] .*

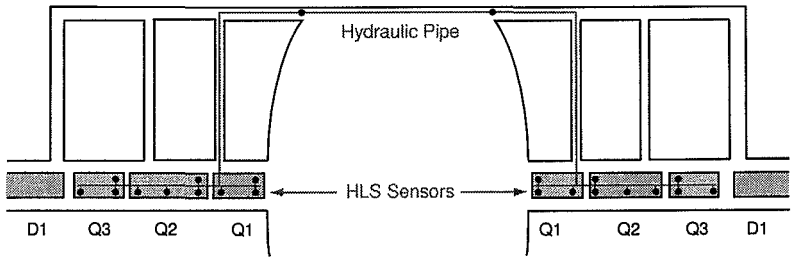


Figure 3.21: *Layout of HLS installation around ATLAS cavern.*

longitudinal position is of lesser interest for the accelerator position but has to be determined for the cavern network by distance measurements.

The Hydrostatic Levelling System works on the principle of communicating vessels, see Figure 3.22. The sensors used at CERN have been manufactured by *Fogale nanotech*, see Figure 3.23. They measure the distance to the liquid surface based on the principle of

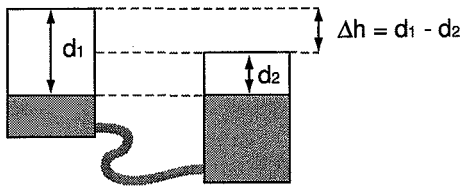


Figure 3.22: *Principle of HLS: Communicating vessels.*

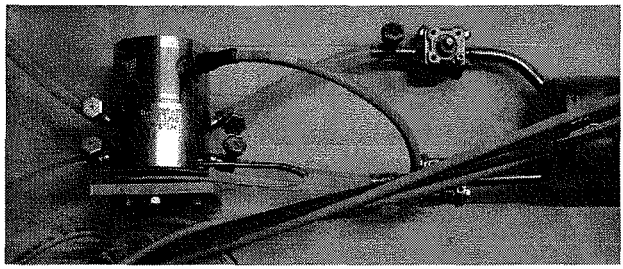


Figure 3.23: *HLS sensor installed in ATLAS survey gallery.*

a capacitor. The accuracy achieved is $< 5\mu$m for a measurement range of ± 2.5 mm, [10]. Effects of pressure and temperature have to be considered as well as tidal effects and gravity anomalies if the distances of the hydraulic network become longer (i.e. > 200 m), [5]. Again the electronics for this system can be remote from the actual measurement system in order to minimize radiation exposure.

WPS and HLS sensors will be installed with mechanical interfaces to each other and to ordinary survey measurement systems by standard reference sockets (described in Section 3.2) mounted on top of the combined WPS/HLS sensor setup. This allows to make the necessary link from the LHC geometry transferred by these systems into the survey galleries and the cavern network.

Unfortunately these systems are not available for the most part of the ATLAS detector installation process. Parts of the HLS and WPS have been installed so far but the important link between the systems as well as the link to LHC geometry are not yet available. As discussed already before in Section 2.4 this link is very important for the ATLAS cavern network as it represents the network datum. This has the effect that the essential connection to the LHC geometry (represented by LHC tunnel reference point monuments until fiducial points on the Inner Triplet Quadrupoles Q_1 Q_2 and Q_3 are installed) is realized by angle and distance measurements passing through the survey galleries which gives a poor network configuration.

Six HLS sensors have been installed on the ATLAS bedplates in order to monitor any movements in this structure during the installation process. These sensors have been linked by geometrical levelling to the cavern network and some sample measurements are included in the results presented in Chapter 6.4.

Digital close range photogrammetry

Digital close range photogrammetry has been extensively used in the context of verification, assembly and fiducialisation of detector parts for all CERN experiments. Likewise many ATLAS detector parts have been measured and verified during their production and assembly at CERN and at external manufacturers. For ATLAS digital photogrammetric measurements have been carried out for some detector systems (e.g. Inner Detector and Tile Calorimeter, see Figure 3.24) asking for high accuracy and measurement of a high number of points.

In the context of the cavern network these measurements are not considered even though some network points are measured for transformation purposes (demanding thus availability of good network point positions). Object point measurements prevail, but this method represents an interesting and very powerful technique in the context of large scale physics experiments by means of its portability and high precision capabilities.

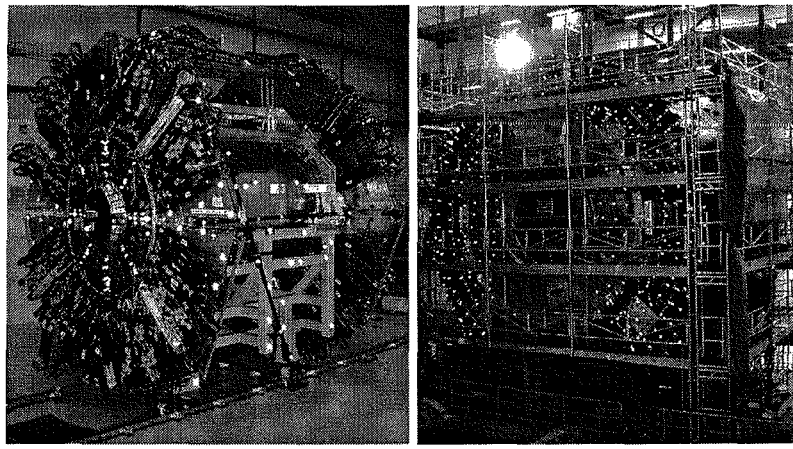


Figure 3.24: *Digital photogrammetry. SCT of the ATLAS Inner Detector (left). Tile Extended Barrel Side A Assembly in the ATLAS cavern (right).*

4

Standard network and deformation analysis

In this chapter standard network and deformation analysis methods are summarized and briefly discussed, considering especially the problem of deforming networks with changing configurations.

The main reference for this chapter is [44]. It is only referred once in each section it applies to. Detailed expressions of most formulas are omitted here and can be found in this reference or any other reference on geodetic network analysis methods. The nomenclature in this chapter follows [16].

4.1 Network analysis

The main aim of network analysis is to assess the quality of a network of geodetic measurements. The definition of a geodetic datum influences the results of network measurements and is also addressed here briefly. The quality of a network is assessed by measures of accuracy and reliability.

In this chapter the formulation of least-squares adjustment according to the Gauss-Markov model is used:

$$\mathbf{y} + \mathbf{e} = \mathbf{A} \cdot \xi \tag{4.1}$$

$$\mathbf{\Sigma_{yy}} = \sigma_0^2 \cdot \mathbf{V} \tag{4.2}$$

where $\mathbf{y}$ is the vector of observations, $\mathbf{e}$ are the residuals, $\mathbf{A}$ is the design matrix representing the functional model of linearized observation equations and ξ is the parameter vector (vector of unknowns). $\mathbf{\Sigma_{yy}}$ as covariance matrix of the observations describes the stochastic model by means of the cofactor-matrix $\mathbf{V}$ scaled by the unit variance σ_0^2. The solution of (4.1) with respect to $\hat{\xi}$, defined as the *best linear uniformly unbiased estimate* (BLUUE) is

$$\hat{\xi} = (\mathbf{A^T P A})^{-1} \mathbf{A^T P y} \tag{4.3}$$

if $\mathbf{A}$ is of full rank and with $\mathbf{P} = \mathbf{V}^{-1}$.

For a detailed discussion of least-squares adjustment and hypothesis testing the reader is referred to [44, 59, 8, 61].

4.1.1 The geodetic datum

A network of survey measurements determines generally only the internal geometry of the concerned network points. In order to derive coordinates in a given reference system the measurements do not give sufficient information. The network is *rank deficient* in terms of the coordinate system definition. The choice of the datum elements relating the internal network geometry to the reference coordinate system directly effects the network's adjustment results. The number of free datum parameters d in a network depends on the dimension of the network and the types of measurements used. It corresponds to the rank deficiency in the design matrix $\mathbf{A}$ describing the relationships between u unknowns by means of the observation equations, $\text{rank}(\mathbf{A}) = u - d$. If only relative observations are available and no datum has been defined, this as well as the matrix of normal equations $\mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}$ are rank deficient, thus they can not be inverted regularly or solved unambiguously. By defining a geodetic datum, one can dispose of the rank deficiency.

A three dimensional network generally has seven free datum parameters: Three translations, three orientations and one scale factor. This is reduced to six if distances have been measured, fixing the scale factor.

If at least two vertical angle measurements are included, the number of free datum parameters is reduced to four, as rotations around X and Y axes are determined. The information inherent in the observations can thus define part of the geodetic datum.

Several methods exist to define the remaining datum parameters, [44]. Generally, two approaches have to be distinguished:

- The datum definition constrains the network's geometry.
- The network is adjusted unconstrained.

Classical (hierarchical) adjustment approaches define a group of points of assumed superior accuracy as datum points, e.g. points that are part of a higher order network. These datum points are fixed and no coordinate changes for these points are allowed in the adjustment. Consequently, no error estimation is available for these points after adjustment. If the number of coordinate components included in this definition is equal to the number of free datum parameters, no constraint is forced on the network, it is adjusted freely. If, however, the number of elements in the datum definition exceeds the number of free datum elements, the network adjustment solution is constrained by this datum definition. Errors in the datum points' relative geometry directly affect the remaining network.

More general methods of defining a datum include conditions on the parameters or additional information in the adjustment model, which dispose of the rank deficiency problem:

- Condition: Minimize parameter estimates for all points: $\|\hat{\xi}\| = \min$. This solution is obtained by applying the algorithm of a *Moore-Penrose matrix inversion* to the matrix of normal equations $\mathbf{N}$. The Moore-Penrose inverse is also referred to as the *pseudo-inverse*. The resulting cofactor matrix $\mathbf{V}_{\hat{\xi}\hat{\xi}}$ has minimum trace $\sum_i \mathbf{V}_{\hat{\xi}\hat{\xi}}(i,i) = \min$. Consequently this method of datum definition is in German referred to as *'Gesamtspurminimierung'*.
- Condition: Minimize parameter estimates for a selected group of datum elements: $\|\hat{\xi}_D\| = \min$. Similarly to the method above the resulting cofactor matrix for the selected datum elements has minimum trace $\sum_i \mathbf{V}_{\hat{\xi}\hat{\xi}_D}(i,i) = \min$, and in German is referred to as *'Teilspurminimierung'*. This method is well adapted for

applications were network points of higher order are used for the datum definition but their accuracy is questionable.

- Additional information: Include stochastic information for datum elements. Datum point coordinates are introduced as additional observation data to the functional adjustment model. The stochastic model is extended by the corresponding stochastic information. This method is well adapted for applications where datum elements with stochastic information exist which is often the case in geodetic networks.

To change the datum definition without re-adjusting the entire model, similarity (S)-transformations can be applied. This method is presented in more detail in Appendix A.

The definition of the geodetic datum is important in deformation analysis problems as the choice of the datum can effect the deformation analysis. Constraints from the datum definition on the network geometry affect the deformation analysis results and have to be avoided. If this is not possible special care needs to be taken to ensure identical datum definitions in all measurement epochs.

4.1.2 Accuracy

The accuracy of a network describes how well the unknown parameters can be determined by the observations available, assuming that the functional and stochastic models are correct. It determines the quality of the network configuration, [44] and is affected by the choice for the datum definition. Several measures of accuracy can be derived from the cofactor matrix of unknowns $\mathbf{V}_{\hat{\xi}\hat{\xi}} = \mathbf{N}^{-1}$. Thus these terms can be derived from the definition of the functional and stochastic models only, no actual observation data is necessary. This makes these terms powerful tools for the planning of measurements.

Accuracy measures are divided into two groups:

- **Local measures of accuracy** refer to parts of the cofactor matrix $\mathbf{V}_{\hat{\xi}\hat{\xi}}$ only. They include standard deviations of individual unknown parameters, confidence-ellipsoids for individual points and relative confidence-ellipsoids between two points.

- **Global measures of accuracy** make use of the full $\mathbf{V}_{\hat{\xi}\hat{\xi}}$ matrix which allows for different network configurations to be compared. Terms include confidence hyper-ellipsoids, measures for homogeneity and isotropy of error ellipsoids and spectral analysis of $\mathbf{V}_{\hat{\xi}\hat{\xi}}$ giving the so called *principal components.* Principal component analysis can be interesting in deformation analysis applications with high parameter correlations, as weak areas in the network can be identified. The worst-determined direction in the network found by principal component analysis, should not coincide with the expected direction of deformation.

Details for all terms can be found in [44].

4.1.3 Reliability

Reliability refers to the ability of the network to control individual measurements and their influence on the unknown parameters. It describes the quality of the network realization. The cofactor-matrix of the residuals $\mathbf{V}_{\hat{\mathbf{e}}\hat{\mathbf{e}}} = \mathbf{V} - \mathbf{A}\mathbf{V}_{\hat{\xi}\hat{\xi}}\mathbf{A}^T$ is used to derive various measures of reliability.

A geodetic network is considered to be reliable, if gross errors in the observation data can be detected (internal reliability) and in case gross errors remain undetected, they have minor influence on the resulting parameter estimation (external reliability).

Redundancy A fundamental term in assessing the reliability of a network is the redundancy number r_i of an individual observation. It is defined as:

$$r_i = \mathbf{R}(i,i) \tag{4.4}$$

with $\mathbf{R}$ being defined in the following way

$$\mathbf{R} = \mathbf{V}_{\hat{\mathbf{e}}\hat{\mathbf{e}}} \cdot \mathbf{P}. \tag{4.5}$$

The redundancy numbers are $0 \leq r_i \leq 1$ for uncorrelated observations. The redundancy R of a network of n observations and u unknown parameters is obtained as

$$R = \text{trace}(\mathbf{R}) = \sum_{i=1}^{n} r_i = n - u. \tag{4.6}$$

The redundancy number of an individual observation describes the contribution of this observation to the overall network redundancy. If it is very small the observation is poorly controlled. Errors in such an observation are difficult to detect but have a major influence on the adjustment results. If the redundancy is large the corresponding observation is very well controlled. Any error in such an observation is easily detectable or if it remains undetected, has a very small influence on the adjustment result. If the redundancy is equal to 1, the observation is completely redundant and thus dispensable.

Gross error detection Tests for gross errors traditionally employ observation residuals $\hat{\mathbf{e}} = \mathbf{A}\hat{\xi} - \mathbf{y}$. If the null hypothesis H_0, stating no gross error being present, is valid

$$H_0 \; : \; E(\mathbf{y}) = \mathbf{A}\xi, \tag{4.7}$$

the residuals $\mathbf{e}$ follow a normal distribution

$$\hat{\mathbf{e}} \sim N\left(\mathbf{0}, \sigma_0^2 \mathbf{V}_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\right). \tag{4.8}$$

Thus a gross error in the observation data can be identified by using the following probability relationship

$$P\{\frac{|\hat{e}_i|}{\sigma_{\hat{e}_i}} > z_{1-\alpha}|H_0\} = 1 - \alpha. \tag{4.9}$$

This simple test is repeated for each observation to identify gross errors at a level of significance of $(1 - \alpha)$.

Internal reliability can be assessed by determining the minimal detectable bias (MDB) in each observation given a certain level of significance $(1 - \alpha)$ and power of the test $(1 - \beta)$. α is referred to as the type I error probability. It is the probability by which an alternative hypothesis H_A is accepted while the null hypothesis H_0 is true. β is referred to as the type II error probability and describes the probability of accepting H_0, if H_A is true in fact. For an illustration of the relationship between α and β see Figure 4.1.

In gross error detection the alternative hypothesis is generally formulated assuming only one gross error Δ_i to be present in the observations, as defined by Baarda in [4]. For every observation the following

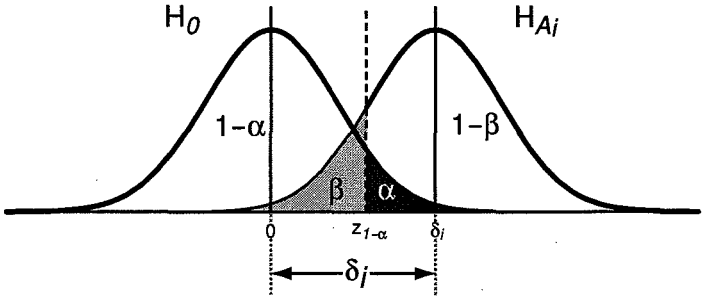


Figure 4.1: *Level of significance* $(1-\alpha)$*, power of the test* $(1-\beta)$ *and non–centrality parameter* δ_i *in the hypothesis test.*

alternative hypothesis can thus be formulated

$$H_{A_i} \; : \; E\left(y_i \,|\, H_{A_i}\right) = E\left(y_i \,|\, H_0\right) + \Delta_i. \tag{4.10}$$

By including the error Δ_i in the functional model Niemeier shows in [44, p. 279] that an estimate $\hat{\Delta}_i$ for the error and its cofactor can be derived

$$\hat{\Delta}_i \;\; = -\frac{\eta_{[i]}\mathbf{P}\hat{\mathbf{e}}}{\eta_{[i]}^T\mathbf{P}\mathbf{V}_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{P}\eta_{[i]}} \tag{4.11}$$

$$v_{\Delta_i} \;\; = \left(\eta_{[i]}^T\mathbf{P}\mathbf{V}_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{P}\eta_{[i]}\right)^{-1} \tag{4.12}$$

where $\eta_{[i]}$ is a unity vector with zeros for all elements but the element i, which is 1. A test statistic Ω_i^2 for this parameter can be derived as

$$\Omega_i^2 = \frac{\hat{\Delta}_i^T\hat{\Delta}_i}{\sigma_0^2 v_{\nabla_i}} = \frac{\hat{\mathbf{e}}^T\mathbf{P}\eta_{[i]}\eta_{[i]}^T\mathbf{P}\hat{\mathbf{e}}}{\sigma_0^2\eta_{[i]}^T\mathbf{P}\mathbf{V}_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{P}\eta_{[i]}} \sim \chi^2_{1-\alpha,1,\delta_i} \tag{4.13}$$

which follows a χ^2–distribution with the non–centrality parameter δ_i. This is due to the fact that Ω_i^2 is the squared form of the variable $\hat{\Delta}_i$ which follows a non-central normal distribution. By executing the test it can be decided if an error in the observation y_i is significant or not. The non–centrality parameter δ_i gives information about the separability between the null and alternative hypothesis and thus the power of the test. It can be determined from the true value $\tilde{\Delta}_i$ of the

gross error or alternatively, as the true error is generally not known, for a given value Δ_i.

$$\delta_i = \frac{\tilde{\Delta}_i^2}{\sigma_0^2 v_{\Delta_i}} = \frac{\tilde{\Delta}_i^2 \eta_{[i]}^T \mathbf{P} \mathbf{V}_{\hat{e}\hat{e}} \mathbf{P} \eta_{[i]}}{\sigma_0^2}. \tag{4.14}$$

In case the observations are uncorrelated ($\mathbf{P} = \mathrm{diag}(p_i)$), the terms described above become simple. Equivalent to (4.4) the redundancy numbers are

$$r_i = (v_{\hat{e}\hat{e}})_{ii} \cdot p_i, \tag{4.15}$$

the estimation for the error and its cofactor in (4.11) become

$$\hat{\Delta}_i = -\frac{\hat{e}_i}{r_i}, \tag{4.16}$$

$$q_{\Delta_i} = -\frac{1}{p_i \cdot r_i}, \tag{4.17}$$

and the non-centrality parameter according to (4.14)

$$\delta_i = \frac{\tilde{\Delta}_i^2 \cdot p_i \cdot ri}{\sigma_0^2}. \tag{4.18}$$

Given a level of significance $(1-\alpha)$ and a chosen value for the power of the test $(1-\beta)$, a non-centrality parameter $\delta_{i,0}$ and consequently a value for the MDB ∇_i can be derived as

$$\nabla_i = \sigma_0 \sqrt{\frac{\delta_{i,0}}{p_i \cdot r_i}}. \tag{4.19}$$

In these simple formulas it can be seen, that the internal reliability is directly affected by the redundancy number of each observation and the observation precision. For a given observations precision, it is easier to detect an error in well controlled observations (e.g. $r_i > 0.6$), as the non-centrality parameter and thus the power of the test increases, see (4.18). The factor r_i emphasizes the error. In a well controlled observation, smaller errors become detectable, according to (4.19).

External reliability describes the influence of an undetected error on the parameter estimates. A single MDB ∇_i in the observation data can change all parameters by $\boldsymbol{\Delta}\hat{\xi}_{\nabla_i}$,

$$\boldsymbol{\Delta}\hat{\xi}_{\nabla_i} = \mathbf{V}_{\hat{\xi}\hat{\xi}}\mathbf{A}\mathbf{P}\eta_{[i]}\nabla_i. \tag{4.20}$$

This term depends on the choice of the datum, as $\boldsymbol{\Delta}\hat{\xi}_{\nabla_i}$ depends on $\mathbf{V}_{\hat{\xi}\hat{\xi}}$. By squaring (4.20) a term θ_i^2 can be defined

$$\begin{aligned}\theta_i^2 &= \boldsymbol{\Delta}\hat{\xi}_{\nabla_i}{}^T\mathbf{V}_{\hat{\xi}\hat{\xi}}{}^{-1}\boldsymbol{\Delta}\hat{\xi}_{\nabla_i} = \nabla_i^2\eta_{[i]}^T\mathbf{P}\mathbf{A}^T\mathbf{V}_{\hat{\xi}\hat{\xi}}\mathbf{V}_{\hat{\xi}\hat{\xi}}{}^{-1}\mathbf{V}_{\hat{\xi}\hat{\xi}}\mathbf{A}\mathbf{P}\eta_{[i]} \\ &= \nabla_i^2\eta_{[i]}^T\mathbf{P}\mathbf{A}^T\mathbf{V}_{\hat{\xi}\hat{\xi}}\mathbf{A}\mathbf{P}\eta_{[i]} = \nabla_i^2\eta_{[i]}^T\mathbf{P}\mathbf{V}_{\hat{y}\hat{y}}\mathbf{P}\eta_{[i]} \\ &= \nabla_i^2\eta_{[i]}^T\mathbf{P}\left(\mathbf{V} - \mathbf{V}_{\hat{e}\hat{e}}\right)\mathbf{P}\eta_{[i]} \\ &= \nabla_i^2\eta_{[i]}^T\mathbf{P}\mathbf{V}\mathbf{P}\eta_{[i]} - \nabla_i^2\eta_{[i]}^T\mathbf{P}\mathbf{V}_{\hat{e}\hat{e}}\mathbf{P}\eta_{[i]}.\end{aligned} \tag{4.21}$$

Assuming uncorrelated observations θ_i^2 follows as

$$\theta_i^2 = \nabla_i^2 \cdot p_i - \nabla_i^2 \cdot p_i \cdot r_i = \nabla_i^2 \cdot p_i\,(1 - r_i)\,. \tag{4.22}$$

This *distortion parameter* θ_i^2 is independent of the datum definition. If the observation redundancy number is larger, the distortion decreases (for same measurement precision and observation error). Thus, as stated before, an error in a well controlled observation has less effect on the network results than a poorer controlled observation.

4.2 Standard deformation analysis

A geodetic network that has been measured in several *epochs* (i.e. at different instances in time) can be analyzed for any changes in the network's geometry. In this section the standard method of congruency testing is shortly reviewed following [44] closely as main reference, but more detailed information can be found in [43, 47, 48].

Changes in the network geometry are detected by a global congruency test. In case a deformation has been detected the reason for it is identified in a subsequent localization procedure.

Information required for each measured epoch i include: the estimated point coordinates $\hat{\xi}_i$, corresponding covariance information $\sigma_{0,i}^2 \cdot \mathbf{V}_{\hat{\xi}_i\hat{\xi}_i}$ and the degree of freedom in the estimation f_i. Generally it is assumed that the network has been measured in identical configuration in each epoch.

Global congruency test Let us assume, a network has been measured in two epochs and needs to be analyzed for possible deformations. The measurements are combined into one model. It is assumed that no correlations exist between the two measurement epochs and that the approximate point coordinates are identical:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}+\begin{bmatrix}\hat{\mathbf{e}}_1\\ \hat{\mathbf{e}}_2\end{bmatrix}=\begin{bmatrix}\mathbf{A}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix}\begin{bmatrix}\hat{\xi}_1\\ \hat{\xi}_2\end{bmatrix},\quad \boldsymbol{\Sigma}_{\mathbf{yy}}=\sigma_0^2\mathbf{V}=\sigma_0^2\begin{bmatrix}\mathbf{V}_{11} & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_{22}\end{bmatrix}. \quad (4.23)$$

The null hypothesis H_0 stating no significant movement between the two epochs can be formulated as

$$H_0 : \quad \hat{\xi}_2-\hat{\xi}_1=\mathbf{0}, \quad (4.24)$$

or alternatively in matrix notation (with unity matrix $\mathbf{I}$)

$$H_0 : \begin{bmatrix}-\mathbf{I} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\hat{\xi}_1\\ \hat{\xi}_2\end{bmatrix}=\mathbf{0}. \quad (4.25)$$

A test term Ω^2 can be formulated as

$$\Omega^2=\left(\hat{\xi}_2-\hat{\xi}_1\right)^T\cdot \left[\begin{bmatrix}-\mathbf{I} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\left(\mathbf{A}_1^T\mathbf{P}_1\mathbf{A}_1\right)^+ & \mathbf{0}\\ \mathbf{0} & \left(\mathbf{A}_2^T\mathbf{P}_2\mathbf{A}_2\right)^+\end{bmatrix}\begin{bmatrix}-\mathbf{I}\\ \mathbf{I}\end{bmatrix}\right]^+\left(\hat{\xi}_2-\hat{\xi}_1\right) \quad (4.26)$$

resulting after multiplication in

$$\Omega^2=\left(\hat{\xi}_2-\hat{\xi}_1\right)^T\left[\left(\mathbf{A}_1^T\mathbf{P}_1\mathbf{A}_1\right)^+ + \left(\mathbf{A}_2^T\mathbf{P}_2\mathbf{A}_2\right)^+\right]^+\left(\hat{\xi}_2-\hat{\xi}_1\right). \quad (4.27)$$

The pseudo-inverse (indicated by $^+$) is used for reasons of generalization in this expression, as singular measurement configurations do not necessarily have to be excluded. (4.27) is further simplified by defining the difference vector ϑ.

$$\vartheta=\hat{\xi}_2-\hat{\xi}_1 \quad (4.28)$$

The corresponding cofactor-matrix $\mathbf{V}_{\vartheta\vartheta}$ follows as

$$\mathbf{V}_{\vartheta\vartheta}=\mathbf{V}_{\hat{\xi}_1\hat{\xi}_1}+\mathbf{V}_{\hat{\xi}_2\hat{\xi}_2}=\left(\mathbf{A}_1^T\mathbf{P}_1\mathbf{A}_1\right)^+ + \left(\mathbf{A}_2^T\mathbf{P}_2\mathbf{A}_2\right)^+. \quad (4.29)$$

It is of rank $h = \text{rank}\,(\mathbf{V}_{\vartheta\vartheta})$ which is in case of identical network configurations and datum definitions for both epochs $h = \text{rank}\left(\mathbf{V}_{\hat{\xi}_1\hat{\xi}_1}\right) = \text{rank}\left(\mathbf{V}_{\hat{\xi}_2\hat{\xi}_2}\right) = u-d$ with d being the number of free datum parameters. Thus (4.27) can be rewritten as

$$\Omega^2 = \vartheta^T \mathbf{V}_{\vartheta\vartheta}{}^{+}\vartheta \tag{4.30}$$

To determine a representative estimate for the value of estimated unit variance $\hat{\sigma}_0^2$ it has to be assured that individual estimates for both epochs $\hat{\sigma}_{0,1}^2 = \frac{\hat{\mathbf{e}}_1^T \mathbf{P}_1 \hat{\mathbf{e}}_1}{f_1}$ and $\hat{\sigma}_{0,2}^2 = \frac{\hat{\mathbf{e}}_2^T \mathbf{P}_2 \hat{\mathbf{e}}_2}{f_2}$ do not differ significantly. The *base* hypothesis is defined as H_B

$$H_B \;:\; E\left(\sigma_{0,1}^2\right) = E\left(\sigma_{0,2}^2\right) \tag{4.31}$$

which is tested by the following relationship

$$P\{\frac{\hat{\sigma}_{0,1}^2}{\hat{\sigma}_{0,2}^2} \leq F_{f_1,f_2,1-\alpha} | H_B\} = 1 - \alpha \tag{4.32}$$

employing the *Fisher* distribution quantile.

If (4.31) is established to be valid the global congruency test can be defined as

$$F = \frac{\vartheta^T \mathbf{V}_{\vartheta\vartheta}{}^{+}\vartheta}{\hat{\sigma}_0^2 \cdot h} = \frac{\Omega^2}{\hat{\sigma}_0^2 \cdot h}, \tag{4.33}$$

with the probability relationship

$$P\{F \leq F_{h,f,1-\alpha} | H_0\} = 1 - \alpha. \tag{4.34}$$

If F exceeds the quantile of the corresponding Fisher-distribution, the null hypothesis H_0 is rejected and deformation has occurred between the two epochs with a level of significance $(1 - \alpha)$. In the following localization procedure it has to be decided which point(s) have moved.

The assumption above on identical network configurations in all epochs is not stringent. Methods employing S-transformations or free adjustments to match identical points can be applied.

Localization In case a significant difference between epochs has been found, points that have moved need to be identified and separated from stable points. The group of stable points forms the reference for deformation values, as the information in the difference vector ϑ is relative. In some applications the reference definition is implied in the problem description, i.e. a group of points is defined as stable or assumed stable because these points are outside the deformation area. Nevertheless such reference assumptions have to be verified in order to avoid any bias in the deformation analysis results by any undetected deformation. In other applications a stable reference has to be identified statistically within the group of all points. In both cases the following iterative test strategy can be applied.

Again the example of only two epochs in identical configurations is used, with p identical points. The global congruency test in (4.34) indicates a significant deformation between the two epochs. Each point is now individually suspected of having moved. In the first step $p-1$ points are assumed to be stable, defining the reference. One point in suspected to have moved. The parameter vector is thus split up in the following way

$$\hat{\xi}_k^j = \begin{bmatrix} \hat{\xi}_{F_k}^j \\ \hat{\xi}_{B_k}^j \end{bmatrix} \quad \text{for } k = 1, \ldots, p. \tag{4.35}$$

$\hat{\xi}_{F_k}^j$ refers to the group of $p-1$ assumed fixed points and $\hat{\xi}_{B_k}^j$ refers to the one point k suspected of having moved. The superscript j refers to the datum definition. It is necessary to make a datum transformation from datum j which includes all identical points to datum i which is represented by the $p-1$ points considered to be fixed. This is achieved by S-transformation, for details see Appendix A.

The difference vector and its cofactor matrix for the fixed points $\hat{\xi}_{F_k}^i$ between the two epochs are defined as, [44, p. 379]

$$\begin{aligned} \vartheta_k &= \left(\hat{\xi}_{F_k}^i\right)_2 - \left(\hat{\xi}_{F_k}^i\right)_1 \\ \vartheta_{\vartheta\vartheta_k} &= \left(\mathbf{V}_{FF_k}^i\right)_1 + \left(\mathbf{V}_{FF_k}^i\right)_2, \end{aligned} \tag{4.36}$$

where the subscript k indicates that the point k is excluded from this difference. Similarly to (4.33) a test term can be defined as

$$F_k = \frac{\vartheta_k^T \mathbf{V}^+_{\vartheta\vartheta_k} \vartheta_k}{\hat{\sigma}_0^2 \cdot h_k}. \tag{4.37}$$

For each of the p points a value for F_k is calculated. The point s corresponding to the smallest value for F_k is considered to be the point which has most probably moved.

If $F_s \leq F_{h_s,f,1-\alpha}$ then s is the only point having moved at the significance level of $(1-\alpha)$. If this is not true, at least one more point has moved. Thus in the next step each of the remaining $p-1$ points is suspected to have moved additionally to point s and each corresponding value F_k is determined. This process is repeated until a set of identical points has been found for which the congruency test is accepted.

This method is also referred to as the *backwards strategy.* In the beginning all points are considered to be stable. Then each of them is checked individually and if tested positively for having moved, is excluded from the group of stable points. In applications with a predefined group of presumably stable reference points these points also need to be checked in this way. Such a process can be followed by a *forward strategy* which checks other identical points considered as moving from the very beginning, for being stable points. In this case one single point is added to the fixed points in each calculation step until no more stable points can be identified.

Applying this standard method of deformation analysis to a large geodetic network poses a complex task for several reasons:

- Many calculations steps are necessary to check for all combinations of points in the backward and possibly also forward strategy.

- If network configurations are different in each epoch, analysis for individual points becomes difficult especially in larger networks. e.g. Point P_1 is measured in epoch 1, 3, 4, 5, 8 and 13. Deformation analysis needs to be carried out for epoch pairs (1, 3), (3, 4), (4, 5), (5, 8) and (8, 13) and possibly all other combination pairs. Point P_2 is measured in epoch 2, 3, 7, 10 and 12. Deformation analysis needs to be carried out for epoch pairs (2, 3), (3, 7), (7, 10) and (10, 12) and possibly all other combination pairs.

- Same datum definitions in each epoch can be achieved by S-transformations or free adjustment. If measurements do not include a link to a predefined 'absolute' reference, it is difficult to relate results to the reference, or to compare results.

4.3 Conclusion

In this chapter some basics of network and deformation analysis have been briefly reviewed. It could be seen that measures of accuracy depend on the measurement configuration and datum definition. Measures of reliability depend on the realization of the network measurements which includes measurement precision but also network configuration and datum definition. It was shown that one of the most important reliability measures in a network is the redundancy of observation data.

Standard deformation analysis methods employ epoch–to–epoch comparison of network measurement results. Clear and thorough analysis of networks in changing configurations and datum definitions is complex and demanding by means of organization of measurement data and results. This is especially true if the size of the network to be analyzed becomes larger. A method which incorporates an easy and structured handling and analysis of the data and the necessary deformation analysis would make this task much more accessible and manageable. A solution to this problem is discussed in the following Chapter 5.

5

An alternative approach to adjust 3D networks subject to deformation and changing configuration

In this chapter a more general approach to geodetic networks is taken, namely as time-variable systems, in contrast to classical methods discussed in Chapter 4.

A short summary of a classification of deformation models is given in Section 5.1. Based on this categorization, a concept to handle these deformation models using a *Kalman Filter* (KF) approach is developed.

The motivation for the use of a KF for geodetic applications in Section 5.2 is followed by Section 5.3 on KF basics. In Section 5.4 the KF equations are set up for a kinematic three dimensional (3D) geodetic network observed in several epochs in changing configurations. Ultimately this KF setup has to be extended to an *adaptive* or *self-tuning*

filter to enable the estimation of the deforming network. The key features of this implementation are presented in Section 5.5. Methods to enhance the filter's performance to identify deformations are discussed, including additional information which is outlined in Section 5.6.

Many terms in the following sections originate from or are strongly related to system or control theory and are explained or interpreted for geodetic matters as they occur. The notation generally found in literature on KF is used in an adapted version. The KF notation diverges from the notation commonly used in geodetic least-squares estimation which is also applied in Chapter 4.

5.1 Deformation models

A commonly agreed classification of deformation models in the geodetic context refers to terms known from system theory. Two main groups of deformation models are differentiated: *Descriptive models* and *cause-response models*, [60].

Descriptive models do not consider any acting forces causing the deformation. If the motion can be described as a function of time a *kinematic model* can be defined. If this is not the case standard deformation analysis methods can be applied based on a *congruence model.*

Cause-response models relate acting forces to resulting deformations. If this relation can be modelled by a time dependent function, a *dynamic model* can be defined. If otherwise the time dependency can not be described or can be neglected, the special case of a *static model* arises.

The dynamic model thus represents the ideal case in terms of analysis capabilities, as it enables to describe the system's behavior best. The definition of a dynamic system requires the relationships between the system's input (i.e. forces) and the output (i.e. deformations) to be well known by quality and quantity, which very often proves to be difficult or even impossible.

5.2 Kalman Filter in geodetic applications

The KF has had a considerable impact in many scientific and engineering branches after its first publication in 1960, [26]. Applications in controls and navigation were among the first to pick up rapidly this new approach and many others followed. For examples of the wide range of applications see [58].

The most common application in the field of geodesy and surveying is the control, navigation and tracking of satellites. A satellite is considered as a moving object whose position is to be determined by measurements from known static ground stations.

Common to all traditional KF applications including navigation and control is the large quantity of repeatedly available data, i.e. data in many different time epochs. In geodetic deformation analysis applications this is generally not the case and data is available in only few epochs.

Pelzer was one of the first to introduce the kinematic approach to handle geodetic, time-variable networks for deformation analysis, see [48, 49].

One of the most thorough and dedicated studies since then was supplied by Heunecke [19], who joins the theory and concepts of system theory, mechanical engineering and engineering geodesy in order to analyze deformation processes. He derived and developed the KF as a special interpretation of least squares adjustment following the Gauss-Helmert and Gauss-Markov models. Heunecke introduced the so called 'Hannover-Filter', an adaptive filter incorporating cause-response deformation models considering the causing forces' parameters as stochastic variables, that can be determined within the filtering process. A very sound theoretical basis to this topic is developed, which does not extend to the practical implementation.

The approach chosen by Heunecke [19] can also be applied to the case of a network like the ATLAS cavern network, but in a modified way. The Hannover filter adapts to deformations acting on the geodetic network by including additional causal physical parameters. For the ATLAS cavern network no cause-response model is available, as discussed in Chapter 2. In such a case the adaptive filter has to be interpreted and solved differently. The approach chosen and described in the following incorporates a descriptive and kinematic deformation model.

A key issue in any KF implementation is the correct identification of the stochastic models in order to avoid divergence and to ensure accurate estimates. The chosen approach iteratively adapts the stochastic models in a robust manner, using statistical tests based on innovation analysis.

The main features of the KF implementation presented in this chapter considering a geodetic network like the ATLAS cavern network can be summarized as follows:

- The geodetic network is interpreted as a *kinematic system.*
- *Changing network configurations* are easily handled in the KF, measurements with *rank deficiency* resulting from insufficient datum definition are possible.
- The main interest is the *correct system description* representing the network. Although deformation analysis in the classical sense is not the main issue it can easily be derived.
- A well known system description can be used in later epochs to *ensure network reliability,* also applicable if the network configuration degrades.
- Accuracy information for each network point is always updated, which can be useful for the planning and management of geodetic measurements referring to the network information.

In the subsequent sections the KF formulation and its theoretical background is described, resulting finally in the detailed description of the implemented adaptive KF approach for the ATLAS cavern network.

5.3 Kalman Filter basics

The KF was first introduced in the commonly known form by R.E. Kalman [26]. The novelty of this paper was a recursive algorithm to solve a time-dynamic problem described in a state–space model. This publication was followed by many others. The main references used here include [6, 15, 21, 35].

The formulation and notation in the following is based on [15]. To avoid continuous repetition, a source is only referenced if different. There are slight changes in the choice of time indexing, which are similarly found in [38]. As a consequence of using KF notation, terms which also occur in geodetic least-squares adjustment might thus be referred to in a different way.

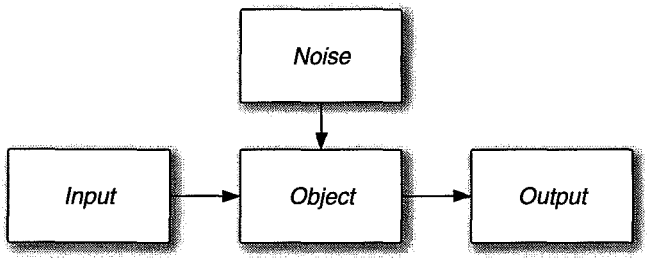


Figure 5.1: *The state–space model.*

For details on random variables, stochastic processes and general least-squares adjustment the reader is referred to [27, 44, 59, 6, 15, 35].

5.3.1 State–space model

Physical, time-variant problems are commonly described by differential equations. For notational and mathematical convenience these equations can be formulated in matrix notation. In system and control theory this formalism is applied in the time domain (in contrast to frequency domain) to yield a so called *state-space model*, see also Figure 5.1.

A linear dynamic system can be described by a first-order differential equation:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t) \cdot \mathbf{x}(t) + \mathbf{G}(t) \cdot \mathbf{w}(t) + \mathbf{L}(t) \cdot \mathbf{u}(t) \tag{5.1}$$

where $\mathbf{x}(t)$ is the system *state vector*, $\mathbf{w}(t)$ is a *random forcing function*, which is assumed to be white noise and describes the system noise, $\mathbf{u}(t)$ is a *deterministic (control) input* (i.e. forces or other causes). $\mathbf{F}(t)$, $\mathbf{G}(t)$ and $\mathbf{L}(t)$ are in general time-dependent, rectangular matrices. The state vector may contain parameters of different nature, for example coordinates, velocities, acceleration, forces, momenta etc. depending on the individual problem at hand. If $\mathbf{u}(t) = 0$, the system is assumed to be not influenced by a deterministic input and the differential equation describes only the kinematics of the system.

Considering only the homogenous part of (5.1) gives

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t) \cdot \mathbf{x}(t). \tag{5.2}$$

This differential equation can be solved if $\mathbf{F}(t)$ is constant (i.e. time invariant) which includes the case where $\mathbf{F}(t)$ is a function of $(t - t_0)$ only. Then the solution $\mathbf{\Phi}(t, t_0)$ is the inverse Laplace transformation $\mathcal{L}^{-1}$ of $(s \cdot \mathbf{I} - \mathbf{F})^{-1}$:

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0) \cdot \mathbf{x}(t_0), \tag{5.3}$$

where $\mathbf{\Phi}(t, t_0)$ represents the *transition matrix* of the system.

Solving for the particular solution of the differential equation (5.1) finally leads to the continuous model description

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0) \cdot \mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau) \cdot \mathbf{G}(\tau) \cdot \mathbf{w}(\tau)\, d\tau + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau) \cdot \mathbf{L}(\tau) \cdot \mathbf{u}(\tau)\, d\tau. \tag{5.4}$$

The continuous model can be transformed into a discrete model:

$$\mathbf{x}_i = \mathbf{\Phi}(i, i-1) \cdot \mathbf{x}_{i-1} + \mathbf{\Gamma}(i-1) \cdot \mathbf{w}_{i-1} + \mathbf{\Lambda}(i-1) \cdot \mathbf{u}_{i-1} \tag{5.5}$$

where

$$\begin{aligned} \mathbf{\Phi}(i, i-1) &= \mathbf{\Phi}(t_i, t_{i-1}) \\ \mathbf{\Gamma}(i-1) \cdot \mathbf{w}_{i-1} &= \int_{t_{i-1}}^{t_i} \mathbf{\Phi}(t_i, \tau) \cdot \mathbf{G}(\tau) \cdot \mathbf{w}(\tau)\, d\tau \\ \mathbf{\Lambda}(i-1) \cdot \mathbf{u}_{i-1} &= \int_{t_{i-1}}^{t_i} \mathbf{\Phi}(t_i, \tau) \cdot \mathbf{L}(\tau) \cdot \mathbf{u}(\tau)\, d\tau. \end{aligned} \tag{5.6}$$

$\mathbf{\Phi}(i, i-1)$ is the transition matrix in time-discrete representation. $\mathbf{w}_{i-1}$ is the (white) system noise process, its relation to the state vector $\mathbf{x}_i$ is described in time-discrete formulation by the corresponding matrix $\mathbf{\Gamma}(i-1)$. The vector $\mathbf{u}_{i-1}$ represents the deterministic input and the rectangular matrix $\mathbf{\Lambda}(i-1)$ summarizes the input's effect in the discrete model.

(5.5) represents a recursive formulation in matrix description of a time-discrete process. Such a *system equation* can then be used to fit a sequence of observed data, thus an estimation problem is defined. The observation data $\mathbf{z}_i$ is introduced by defining the *measurement equation* which relates the observed data to the state vector:

$$\mathbf{z}_i = \mathbf{H}_i \cdot \mathbf{x}_i - \mathbf{v}_i, \tag{5.7}$$

where $\mathbf{v}_i$ is the measurement noise, which is as the system noise before assumed to be white noise. $\mathbf{H}_i$ is the matrix of linear or linearized relationships between the state vector and the measurements.

Here, a time-discrete representation of the measurement equation has directly been chosen, which is most suitable for sampled observation data.

As indicated before the measurement noise $\mathbf{v}_i$ and the system noise $\mathbf{w}_{i-1}$ in (5.5) are to be understood as white-gaussian sequences that are uncorrelated.

$$\begin{aligned} &E\{\mathbf{w}_{i-1}\} = 0; && E\{\mathbf{w}_{i-1} \cdot \mathbf{w}_{j-1}^T\} = \mathbf{Q}_{i-1} \cdot \delta_{ij} \\ &E\{\mathbf{v}_i\} = 0; && E\{\mathbf{v}_i \cdot \mathbf{v}_j^T\} = \mathbf{R}_i \cdot \delta_{ij} \\ &E\{\mathbf{v}_i \cdot \mathbf{w}_{j-1}^T\} = 0, && \text{for all } i, j, \end{aligned} \tag{5.8}$$

δ_{ij} being the *Kronecker-Delta* function defined as $\delta_{ij} = 1$, if $i = j$ and $\delta_{ij} = 0$ otherwise. $\mathbf{R}_i, \mathbf{Q}_{i-1}$ are the covariance matrices of the measurement and system noise, respectively.

The measurement and system noise describe model disturbances and noise corruption that affect the system, but also uncertainty about the model.

5.3.2 Discrete Kalman Filter

The discrete KF determines a minimum mean-square error (MMSE) estimate for the state vector as described in (5.5) based on measurement data included in (5.7).

The proof of the KF properties that lead to a MMSE estimate is given in Kalman's original paper [26]. Kalman's original derivation is based on the Orthogonal Projection Lemma which states, that the minimum estimate is the orthogonal projection of the state on the measurement space. Other derivations can be found in the literature and

include different approaches (e.g. [35, 21] taking conditional probability approaches). In [19] the KF is derived as special cases of the Gauss-Helmert and Gauss-Markov models providing additional possibilities for statistical interpretation interesting to geodetic applications which are missing in other formulations. An extended innovation analysis enabled by this approach is described in the later Section 5.5.2. Other useful terms include the terms of *observability*, *controllability* and *disturbability* summarized in Section 5.3.5.

The estimate of the state vector $\hat{\mathbf{x}}_{i|i}$ at a time t_i is obtained as an *update* of the previous state vector by incorporating information obtained by adding the measurements $\mathbf{z_i}$:

$$\hat{\mathbf{x}}_{i|i} = \hat{\mathbf{x}}_{i|i-1} + \mathbf{K}_i(\mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}), \tag{5.9}$$

where $\hat{\mathbf{x}}_{i|i}$ represents the estimate for the state at time i based on measurement information up to and including time i, $\mathbf{z}_i$. The vector $\hat{\mathbf{x}}_{i|i-1}$ represents the estimate (*prediction*) for the state at time i based on measurement information up to and including time i, $\mathbf{z}_{i-1}$. The term $\mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}$ describes the expected measurement values for time i, based on the measurement equations and the state, predicted for time i. The *gain matrix* $\mathbf{K}_i$ links the real and predicted measurement values to the updated state vector, by means of the mean-square error minimum.

Similarly, the update of the state vector's covariance matrix $\mathbf{P}_{i|i}$ is obtained as

$$\mathbf{P}_{i|i} = (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{P}_{i|i-1}. \tag{5.10}$$

An alternative, numerically more stable form, referred to as the *Joseph form* is defined as, [6, p. 261]:

$$\mathbf{P}_{i|i} = (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{P}_{i|i-1}(\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)^T + \mathbf{K}_i\mathbf{R}_i\mathbf{K}_i^T. \tag{5.11}$$

Its symmetry reduces some numerical problems due to e.g. large initialization values for the state uncertainties.

In (5.9) $\hat{\mathbf{x}}_{i|i-1}$ is the state estimate predicted for time i, based on measurement data up to and including time $i-1$, $\mathbf{z}_{i-1}$. It is

$$\hat{\mathbf{x}}_{i|i-1} = \mathbf{\Phi}(i, i-1)\hat{\mathbf{x}}_{i-1|i-1} + \mathbf{\Gamma}(i-1)\mathbf{w}_{i-1}. \tag{5.12}$$

Because of

$$E\{\mathbf{w}_{i-1}\} = \mathbf{0}, \quad E\{\mathbf{w}_{i-1}\mathbf{w}_{i-1}^T\} = \mathbf{Q}_{i-1} \tag{5.13}$$

the prediction follows as:

$$\hat{\mathbf{x}}_{i|i-1} = \mathbf{\Phi}(i, i-1)\, \hat{\mathbf{x}}_{i-1|i-1} \tag{5.14}$$

$$\mathbf{P}_{i|i-1} = \mathbf{\Phi}(i, i-1)\mathbf{P}_{i-1|i-1}\mathbf{\Phi}^T(i, i-1) + \mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1), \tag{5.15}$$

where $\mathbf{\Phi}(i, i-1)$ is the *transition matrix* which propagates the state variables from time $i-1$ to i. Some of its significant properties include, [15, p. 58f]:

- $\mathbf{\Phi}(i, i) = \mathbf{I}$
- for any i and $i-2$:

$$\mathbf{\Phi}(i, i-2) = \mathbf{\Phi}(i, i-1) \cdot \mathbf{\Phi}(i-1, i-2) \tag{5.16}$$

- non-singularity:

$$\mathbf{\Phi}(i, i-1) \cdot \mathbf{\Phi}(i, i-1) = \mathbf{\Phi}(i, i) = \mathbf{I} \tag{5.17}$$

premultiplying this expression by $\mathbf{\Phi}^{-1}(i, i-1)$ yields

$$\mathbf{\Phi}^{-1}(i, i-1) = \mathbf{\Phi}(i-1, i). \tag{5.18}$$

As mentioned before, $\mathbf{K}_i$ as in (5.9), is referred to as the *Kalman gain* matrix as it specifies the new observations' influence on the state vector update. If the gain is small, then the new estimate largely relies on the previous state vector as more weight is given to the system than on the new measurements. The gain matrix is obtained by

$$\mathbf{K}_i = \mathbf{P}_{i|i-1}\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i)^{-1}, \tag{5.19}$$

where $\mathbf{P}_{i|i-1}$ is the predicted error covariance matrix for the state at time i.

To start the filter at time $t = 0$, initial estimates for $\mathbf{x}$ and $\mathbf{P}$ have to be provided. At this point it is sufficient to define the initial state as a random vector

$$\begin{aligned} &E\{\mathbf{x}_0\} = \hat{\mathbf{x}}_0; \\ &E\{(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T\} = \mathbf{P}_{0|0}. \end{aligned} \tag{5.20}$$

The term $(\mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1})$ in (5.9) is also referred to as the vector of *innovation* $\mathbf{d}$. The corresponding covariance matrix is $\mathbf{D}$. The definitions are summarized as

$$\mathbf{d}_i = \mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}, \tag{5.21}$$

$$\mathbf{D}_i = \mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i. \tag{5.22}$$

The innovations represent the discrepancy between reality (measurement equation) and the model (system equation). Thus it can be used to determine the suitability of a certain model to represent specific data. If $\mathbf{K}_i$ is the optimal gain for the process described, the innovations process $\mathbf{d}$ is a white noise process. This innovation property is a key performance indicator of the KF and is pursued later in more detail, see Section 5.3.4 and 5.5.2.
A summary of the algorithm, its inputs and variables, is depicted in Figure 5.2.
A remark on notation: Gelb [15] denotes the innovation as ν, and does not list a specific term for its covariance matrix. Maybeck in [35] prefers the term residuals $\mathbf{r}_i$. As residual is already a heavily used term in classical geodetic least-squares calculations, the term innovation is preferable. Nevertheless it should be noted here that some methods summarized by 'residual monitoring' make use of the definition as by Maybeck. Residual monitoring in this context corresponds to later discussed innovation analysis methods, forming the basis for adaptive filtering. Many other authors use $\mathbf{d}$ for the innovations and $\mathbf{D}$ for its covariance. These terms are also used here.

5.3.3 Nonlinear filtering

The KF algorithm as described in the previous Section 5.3.2 is formulated in linear matrix algebra. A major part of real, physical problems are nonlinear. Still, the KF can be effectively used for nonlinear problems. The nonlinearity can occur in the description of the system dynamics or in the measurement equations, or both. The system equation thus reads

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{G}(t) \cdot \mathbf{w}(t) + \mathbf{L}(t) \cdot \mathbf{u}(t), \tag{5.23}$$

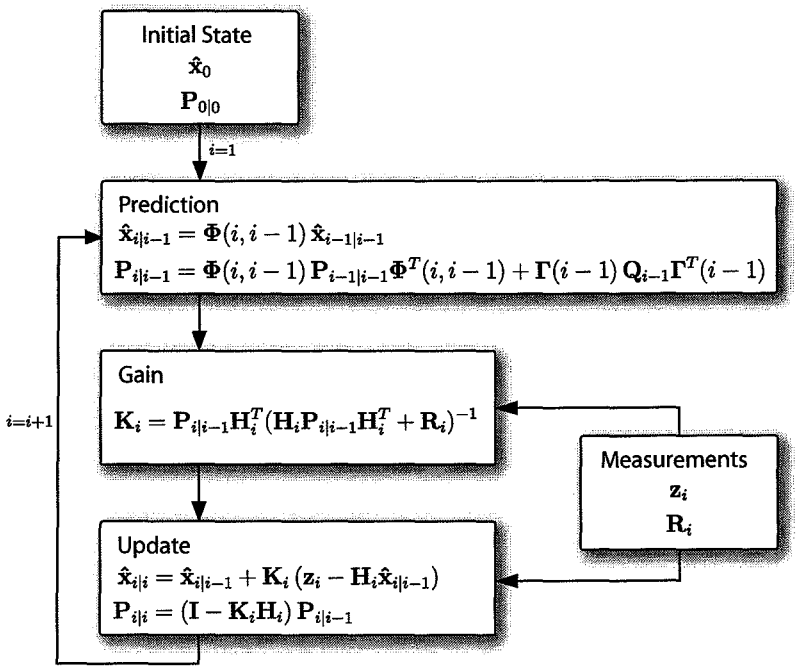


Figure 5.2: *The discrete Kalman Filter.*

where $\mathbf{f}$ is a nonlinear function of the state. The matrices describing the forcing functions $\mathbf{G}(t)$ and $\mathbf{L}(t)$ are assumed to be linear.

Similarly the measurement equations are obtained

$$\mathbf{z} = \mathbf{h}(\mathbf{x}(t)) - \mathbf{v}, \tag{5.24}$$

where $\mathbf{h}$ is a nonlinear measurement equation.

It is possible to linearize around nominal values if an approximation for the state variables is available (i.e. a trajectory in the state space model). If a nominal trajectory is given, which is independent of the measurements included in the filter, the resulting algorithm is referred to as *linearized KF*. If, on the other hand, linearization occurs around the most recently updated state estimate, that has been determined from measurement data, the resulting filter is called an *extended KF*.

If $\mathbf{x}^*(t)$ refers to the nominal trajectory, the state variable definition, (5.23) and (5.24) become

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}^*(t) + \Delta\mathbf{x}(t); \\ \dot{\mathbf{x}}^* + \Delta\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}^* + \Delta\mathbf{x}, t) + \mathbf{G}(t) \cdot \mathbf{w}(t) + \mathbf{L}(t) \cdot \mathbf{u}(t); \\ \mathbf{z} &= \mathbf{h}(\mathbf{x}^* + \Delta\mathbf{x}, t) - \mathbf{v}(t). \end{aligned} \tag{5.25}$$

Assuming that the correction term $\Delta\mathbf{x}$ is small, a Taylor series expansion of the nonlinear equations to the first order is carried out, resulting in the following linear approximations

$$\begin{aligned} \dot{\mathbf{x}}^* + \Delta\dot{\mathbf{x}} &\approx \mathbf{f}(\mathbf{x}^*, t) + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}*} \cdot \Delta\mathbf{x} + \mathbf{G}(t) \cdot \mathbf{w}(t) + \mathbf{L}(t) \cdot \mathbf{u}(t); \\ \mathbf{z} &\approx \mathbf{h}(\mathbf{x}^*, t) + \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}*} \cdot \Delta\mathbf{x} - \mathbf{v}(t), \end{aligned} \tag{5.26}$$

where the matrices of partial derivatives $\left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right]$ and $\left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right]$ are the equivalents to the matrices $\mathbf{F}$ and $\mathbf{H}$ from Section 5.3.1 respectively:

$$\mathbf{F} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{f}_1}{\partial \mathbf{x}1} & \frac{\partial \mathbf{f}_1}{\partial \mathbf{x}2} & \cdots \\ \frac{\partial \mathbf{f}_2}{\partial \mathbf{x}1} & \frac{\partial \mathbf{f}_2}{\partial \mathbf{x}2} & \cdots \\ \vdots & & \end{bmatrix}; \quad \mathbf{H} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{h}_1}{\partial \mathbf{x}1} & \frac{\partial \mathbf{h}_1}{\partial \mathbf{x}2} & \cdots \\ \frac{\partial \mathbf{h}_2}{\partial \mathbf{x}1} & \frac{\partial \mathbf{h}_2}{\partial \mathbf{x}2} & \cdots \\ \vdots & & \end{bmatrix}. \tag{5.27}$$

If the nominal trajectory $\mathbf{x}^*(t)$ is chosen to satisfy the differential equation $\dot{\mathbf{x}}^* = \mathbf{f}(\mathbf{x}^*, t)$, the linearized model reduces to an incremental model where only the increment of the state $\Delta\mathbf{x}$ is the unknown to be estimated:

$$\begin{aligned} \Delta\dot{\mathbf{x}} &= \mathbf{F} \cdot \Delta\mathbf{x} + \mathbf{G}(t) \cdot \mathbf{w}(t) + \mathbf{L}(t) \cdot \mathbf{u}(t); \\ \mathbf{z} - \mathbf{h}(\mathbf{x}^*, t) &= \mathbf{H} \cdot \Delta\mathbf{x} - \mathbf{v}(t). \end{aligned} \tag{5.28}$$

For the remaining incremental state variables $\Delta\mathbf{x}$ the update equations in the discrete case (compare (5.9)) become

$$\Delta\hat{\mathbf{x}}_{i|i} = \Delta\hat{\mathbf{x}}_{i|i-1} + \mathbf{K}_i(\mathbf{z}_i - \mathbf{h}(\mathbf{x}_i^*) - \mathbf{H}_i\Delta\hat{\mathbf{x}}_{i|i-1}), \tag{5.29}$$

where $\mathbf{z}_i - \mathbf{h}(\mathbf{x}_i^*)$ is the incremental measurement (compare (5.28)). For the extended KF, where the 'best' state estimate available is used as nominal trajectory, the 'corrective' terms $\mathbf{h}(\mathbf{x}_i^*)$ and $\mathbf{H}_i\Delta\hat{\mathbf{x}}_{i|i-1}$ are summarized to give a predicted measurement $\hat{\mathbf{z}}_{i|i-1} = \mathbf{h}(\mathbf{x}_i^*) + \mathbf{H}_i\Delta\hat{\mathbf{x}}_{i|i-1}$. Then the measurement residuals, or innovations, become

$$\mathbf{d}_i = \mathbf{z}_i - \hat{\mathbf{z}}_{i|i-1}. \tag{5.30}$$

Brown and Hwang show in [6] that it is easy with this method to keep track of the absolute estimates in the filter. Considering (5.30), (5.29) can be extended again by the nominal trajectory $\mathbf{x}_i^*$ to give

$$\mathbf{x}_i^* + \Delta\hat{\mathbf{x}}_{i|i} = \mathbf{x}_i^* + \Delta\hat{\mathbf{x}}_{i|i-1} + \mathbf{K}_i(\mathbf{z}_i - \hat{\mathbf{z}}_{i|i-1}) \tag{5.31}$$

which is equivalent to the original form (5.9), as $\mathbf{x}_i^* = \hat{\mathbf{x}}_{i|i-1}$. In the case of the extended KF the incremental $\Delta\hat{\mathbf{x}}_{i|i}$ after the update of the absolute state variables reduces to zero, thus its projection to the next epoch $i+1$ vanishes. Thus $\hat{\mathbf{z}}_{i+1|i}$ reduces to $\hat{\mathbf{z}}_{i+1|i} = \mathbf{h}(\hat{\mathbf{x}}_{i+1|i})$.

5.3.4 Innovation Analysis - Adaptive filtering

As noted before in Section 5.3.2 the vector of innovations $\mathbf{d}$, or simply referred to as the innovation, is a key term in the KF algorithm. It describes the discrepancy between the system description and the measurement information. In this section the properties of $\mathbf{d}$ are briefly presented which can be used to correctly identify the stochastic models in the KF. The correct stochastic models are essential in any KF implementation for the algorithm to work and to avoid divergence. A KF implementation based on the analysis of the innovations in this context is generally referred to as an adaptive KF and possibilities to derive such an algorithm from the innovation properties are discussed in this section and are later applied in Section 5.5.2.

The vector of innovations is defined as

$$\mathbf{d}_i = \mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}. \tag{5.21}$$

The innovations represent the part of information in $\mathbf{z}_i$, that is new to the system knowledge $\mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}$. The latter describes, what the observation data are expected to be, given the knowledge and measurements of

the system up to and including time $i-1$, thus it can also be interpreted as

$$\mathbf{d}_i = \mathbf{z}_i - E\{\mathbf{z}_i|\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{i-1}\}. \tag{5.32}$$

Maybeck derives in [35, p. 228] that $\mathbf{d}_i$ is independent of $\{\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{i-1}\}$ and therefore also independent of any $\mathbf{d}_k$, for $k < i$. Accordingly, the *innovation property*, [25], states:

$$E\{\mathbf{d}_i \cdot \mathbf{d}_j^T\} = 0 \qquad \text{for} \quad i \neq j. \tag{5.33}$$

The innovation process is referred to as a *white process*. Its mean and covariance are

$$\begin{aligned} E\{\mathbf{d}_i\} &= 0 \\ E\{\mathbf{d}_i \cdot \mathbf{d}_i^T\} &= \mathbf{H}_i \mathbf{P}_{i|i-1} \mathbf{H}_i^T + \mathbf{R}_i = \mathbf{D}_i. \end{aligned} \tag{5.34}$$

It can be shown (see, e.g. [35, p. 228f] or [1, p. 100f]) that as $\mathbf{d}_i$ is a linear combination of jointly Gaussian random variables, it is Gaussian for all i.

The information contained in the innovation $\mathbf{d}_i$ describes the discrepancy between the actual measurement data and the best measurement predictions based on the filter model. This is an indicator for erroneous model formulation, which might also be caused by changes in the real system. The steady-state of the filter and the *optimal* gain is found, when no information is contained in the innovations. Optimal is defined in the sense of the MMSE, resulting in the best state estimate considering all measurement data and previous knowledge about the system. The optimal filter can hardly be achieved in real applications and very often the a priori stochastic noise model can only be chosen approximately, resulting in a suboptimal filter. With more data available to the filter the stochastic information needs to be readjusted based on information obtained from measurement data. This task is often referred to as *filter tuning* and the corresponding algorithms as *adaptive* or *self-tuning estimation* algorithms.

The term *adaptive* filter is a heavily used term and a variety of different approaches are named by it. Common to all is that some kind of uncertainty within the algorithm is to be minimized by analyzing the

innovations of the filtering process to adjust the system model with respect to the real system. This includes methods to extend the state vector by more parameters (e.g. physical parameters, [19]) or approaches to identify uncertainties in various system model formulation terms. This section focuses as already outlined briefly, on uncertainties in the stochastic model, represented by noise variance terms $\mathbf{Q}_{i-1}$ related to the system noise vector $\mathbf{w}_{i-1}$ and $\mathbf{R}_i$ related to the measurement noise vector $\mathbf{v}_i$.

Various methods of exploitation of the innovation property are proposed in the literature to find the correct noise variance matrix. Mehra gives a summary of methods and groups them as follows, [39]:

- *Maximum likelihood estimation:* Discussed in great detail in [36]. This approach assumes constant unknown parameters, or at least constant-over-N-steps.

- *Bayesian and multiple model estimation:* in e.g. [6]. Different possible realizations of unknown system parameters ($\mathbf{Q}_{i-1}$ or $\mathbf{R}_i$ elements) are analyzed. One set of parameters is considered to relate to a particular KF implementation. The optimal estimate is found as the weighted sum of all KF implementations, where weights are defined by probability density functions of system parameters conditioned on observations. The parameters are assumed to be constant over time.

- *Correlation methods or residual 'whitening':* These methods are based on correlation analysis of sequences of innovations, [38].

- *Covariance matching techniques:* The unknown variable(s) $\mathbf{Q}_{i-1}$ or $\mathbf{R}_i$ are modified until the innovation covariances match the theoretical values, [21].

These methods all rely on the fact that the unknown values in $\mathbf{Q}_{i-1}$ and $\mathbf{R}_i$ are constant over time. Time series of innovation sequences are necessary to do this, making use of series of measurement values of the same nature throughout the whole process. As is seen later in Chapter 5.5, both issues are unsatisfiable conditions for applications like a geodetic network with changing configurations.

A less demanding and easily implemented method to adapt the stochastic system noise model is *stochastic stabilization* (also found in

the literature as **Q**-Stabilization) as proposed by Schrick, [57]. This method was also suggested by Pelzer, [49] in the context of deformation analysis for kinematic geodetic networks. Only the system noise variance $\mathbf{Q}_{i-1}$ is assumed unknown. Its difficult direct derivation is compensated by an iterative increase of the noise variance (or variance inflation) until the system has sufficiently adapted to incorporate innovation information, which 'stabilizes' the filter. One condition for this method to work is that the system state is *completely disturbable*, a term which is defined in the following Section 5.3.5. Disadvantages of iterative noise variance augmentation are outlined in [19]: The noise uncertainty level is unnaturally high and remains so until the next measurement information is available, resulting in a suboptimal filter behavior. It needs to be addressed for a particular application if this is indeed wanted or an acceptable behavior.

It should be noted that interventions into the stochastic model in a filtering process causes (deliberate) changes in the very important relation between the system noise stochastics and measurement stochastics. So, any conclusion on the stochastic system model is based on the assumption of a correct measurement stochastic model. Any undetected gross errors in the measurement data bias the filter. Later in this chapter it is shown that an adaptive filter can incorporate also gross error detection in the observation data in parallel to identifying the correct system noise stochastic model.

5.3.5 Some characteristics of the KF

In order to better understand and illustrate some processes in the KF, some characteristic terms are introduced. The details on the structure and importance of these terms is kept to a minimum here, becoming apparent in the application to the practical example of the kinematic network in later sections.

If the KF is applied to a changing measurement configuration, it is important to keep track of the availability of information to estimate the state variables. In the context of stochastic stabilization, it has to be assured that the system description can be 'disturbed' (or excited) by changes in the stochastic system noise model. Similarly, if a deterministic input **u** applies, it is of interest, if this input has an effect on the state. Three terms are defined to analyze the structure of the KF, [15, 19]: *Observability, controllability* and *disturbability.*

Observability defines the ability to determine the state variables from available measurements. The system description (5.14) reads:

$$\hat{\mathbf{x}}_{i|i-1} = \mathbf{\Phi}(i, i-1) \cdot \hat{\mathbf{x}}_{i-1|i-1}. \tag{5.35}$$

It is assumed that noise-free measurements are available:

$$\mathbf{z}_i = \mathbf{H}_i \cdot \mathbf{x}_i. \tag{5.36}$$

Summarized for all, say n, epochs the two above expressions can be combined to give

$$\begin{aligned} \mathbf{z}_0 &= \mathbf{H}_0 \cdot \mathbf{x}_0 \\ \mathbf{z}_1 &= \mathbf{H}_1 \cdot \mathbf{x}_1 \quad = \mathbf{H}_1 \cdot \mathbf{\Phi}(1,0) \cdot \mathbf{x}_0 \\ &\vdots \qquad\qquad \vdots \\ \mathbf{z}_n &= \mathbf{H}_n \cdot \mathbf{x}_n \quad = \mathbf{H}_n \cdot \mathbf{\Phi}(n,0) \cdot \mathbf{x}_0 \end{aligned} \tag{5.37}$$

or in matrix notation

$$\begin{bmatrix} \mathbf{z}_0 \\ \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} - \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \cdot \mathbf{\Phi}(1,0) \\ \vdots \\ \mathbf{H}_n \cdot \mathbf{\Phi}(n,0) \end{bmatrix} \cdot \mathbf{x}_0. \tag{5.38}$$

The matrix product term above can be summarized as matrix $\mathbf{\Xi}^T$.

$$\mathbf{\Xi} = \left[\mathbf{H}_0^T \;\middle|\; \mathbf{\Phi}^T(1,0) \cdot \mathbf{H}_1^T \;\middle|\; \cdots \;\middle|\; \mathbf{\Phi}^T(n,0) \cdot \mathbf{H}_n^T\right]. \tag{5.39}$$

In order to determine values for the state vector at time 0, $\mathbf{x}_0$ and at any other time i, $\mathbf{x}_i$, the matrix $\mathbf{\Xi}^T$ must be invertible, i.e. non-singular. Then the system is fully *observable*. To fulfill this condition of observability, $\mathbf{\Xi}$ must have full rank (i.e. $\text{rank}(\mathbf{\Xi}) = r = u$ where u is the dimension of the state vector $\mathbf{x}$). The symmetric matrix product $(\mathbf{\Xi}^T \cdot \mathbf{\Xi})$ is referred to as the *observability matrix*, [19].

The structure of the matrix reveals, that if the individual matrix products $\mathbf{\Phi}^T(i,0) \cdot \mathbf{H}_i^T$ are not of full rank, the composite matrix $\mathbf{\Xi}$ can still be of full rank. Linear dependent columns in $\mathbf{\Phi}^T(i,0) \cdot \mathbf{H}_i^T$ are added to columns in $\mathbf{\Xi}$ where this composition may cause the rows of the complete matrix to be independent from each other. Thus state vector

entries can be observable even if no direct observations of these states are available, e.g. point coordinate velocities. This is mainly due to the properties of the transition matrix, as can be seen later in Section 5.5.4 where the observability term is derived for the example of a kinematic network.

In case the rank $r = \text{rank}(\boldsymbol{\Xi}) < u$, only r state variables can be observed or determined by the observational data.

Controllability Controllability determines under which conditions the state vector can be completely controlled by a deterministic input in systems that are subject to control inputs ($\mathbf{u}_{i-1} \neq \mathbf{0}$). According to [19] the matrix $\boldsymbol{\Theta}^T$ is defined as:

$$\boldsymbol{\Theta}^T = \left[\boldsymbol{\Lambda} \quad | \quad \boldsymbol{\Phi}(1,0)\cdot\boldsymbol{\Lambda} \quad | \quad \cdots \quad | \quad \boldsymbol{\Phi}(n,0)\cdot\boldsymbol{\Lambda}\right]. \tag{5.40}$$

Full controllability which implies that the state vector can take any desired value, is achieved if the matrix $\boldsymbol{\Theta}^T$ has full rank. The *controllability matrix* is defined as $(\boldsymbol{\Theta}^T \cdot \boldsymbol{\Theta})$, [19].

Disturbability describes the ability of a system to activate state variables by system noise disturbance $\mathbf{w}_{i-1}$ and the corresponding covariance matrix $\mathbf{Q}_{i-1}$. For the stochastic stabilization outlined in Section 5.3.4 it is essential that the system is fully disturbable, implying that changes in the system noise error affect the state vector.

$$\boldsymbol{\Psi}^T = \left[\boldsymbol{\Gamma}(1) \quad | \quad \boldsymbol{\Phi}(2,1)\cdot\boldsymbol{\Gamma}(1) \quad | \quad \cdots \quad | \quad \boldsymbol{\Phi}(n,1)\cdot\boldsymbol{\Gamma}(1)\right]. \tag{5.41}$$

Full disturbability is obtained if matrix $\boldsymbol{\Psi}^T$ has full rank. The *disturbability matrix* is defined as $(\boldsymbol{\Psi}^T \cdot \boldsymbol{\Psi})$, [19]. It is derived for the example of a kinematic network in Section 5.5.4.

5.4 The Kalman Filter equations for a kinematic 3D geodetic network

In this section I formulate all necessary equations to setup the KF for a kinematic 3D geodetic network. The state vector $\mathbf{x}$ is composed of components for position, velocity and acceleration for each point. Kinematic system equations for an unforced, uniformly accelerated motion

are set up and the necessary terms derived, including the transition matrix $\mathbf{\Phi}(i, i-1)$ and the noise matrix $\mathbf{\Gamma}(i-1)$.

5.4.1 From equations of motion to system equations

The physical, time-variant problem description expressed in differential equations is formulated in a state–space model. As defined in (5.1) a linear, dynamic problem in general is described in matrix notation as

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t) \cdot \mathbf{x}(t) + \mathbf{G}(t) \cdot \mathbf{w}(t) + \mathbf{L}(t) \cdot \mathbf{u}(t). \tag{5.1}$$

A deforming, three dimensional geodetic network, with little knowledge of the expected movements of the network points, is modelled by a system equation for an unforced ($\mathbf{u}(t) = 0$), uniformly accelerated motion. This implies that jerk (third derivation of position with respect to time) and higher order terms of motion are considered as system noise and are thus not explicitly modelled:

$$\dddot{\mathbf{y}}(t) = \mathbf{w}(t). \tag{5.42}$$

With the following definitions

$$\begin{aligned} \mathbf{x_1}(t) &= \mathbf{y} && \text{position,} \\ \mathbf{x_2}(t) &= \dot{\mathbf{y}} && \text{velocity,} \\ \mathbf{x_3}(t) &= \ddot{\mathbf{y}} && \text{acceleration,} \end{aligned} \tag{5.43}$$

a system of differential equations is defined as

$$\begin{aligned} \dot{\mathbf{x}}_1(t) &= \dot{\mathbf{y}} = \mathbf{x_2}(t) = \mathbf{0} \cdot \mathbf{x_1}(t) + \mathbf{I} \cdot \mathbf{x_2}(t) + \mathbf{0} \cdot \mathbf{x_3}(t) + \mathbf{0} \cdot \mathbf{w}(t) \\ \dot{\mathbf{x}}_2(t) &= \ddot{\mathbf{y}} = \mathbf{x_3}(t) = \mathbf{0} \cdot \mathbf{x_1}(t) + \mathbf{0} \cdot \mathbf{x_2}(t) + \mathbf{I} \cdot \mathbf{x_3}(t) + \mathbf{0} \cdot \mathbf{w}(t) \\ \dot{\mathbf{x}}_3(t) &= \dddot{\mathbf{y}} = \mathbf{w}(t) = \mathbf{0} \cdot \mathbf{x_1}(t) + \mathbf{0} \cdot \mathbf{x_2}(t) + \mathbf{0} \cdot \mathbf{x_3}(t) + \mathbf{I} \cdot \mathbf{w}(t). \end{aligned} \tag{5.44}$$

In matrix notation this gives the system differential equation:

$$\begin{aligned} &\dot{\mathbf{x}}(t) = \mathbf{F} \cdot \mathbf{x}(t) + \mathbf{G} \cdot \mathbf{w}(t), \\ &\text{with } \mathbf{F} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathbf{G} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix}. \end{aligned} \tag{5.45}$$

The homogenous solution to the system differential equation is found (compare with (5.3)) as the solution to

$$\dot{\mathbf{x}}(t) = \mathbf{F} \cdot \mathbf{x}(\mathbf{t}). \tag{5.2}$$

Splitting $\dot{\mathbf{x}}(t)$ up into its three components again gives

$$\begin{aligned} \dot{\mathbf{x_1}}(t) &= \mathbf{x_2}(t), \\ \dot{\mathbf{x_2}}(t) &= \mathbf{x_3}(t), \\ \dot{\mathbf{x_3}}(t) &= \mathbf{0}. \end{aligned} \tag{5.46}$$

Integration over $\int_{t_0}^{t}$ with $(\Delta t = t - t_0)$), and back-substitution gives

$$\begin{array}{lcllllll} \mathbf{x_1}(t) & = & \mathbf{I} \cdot \mathbf{c_1} & + \Delta t \cdot & \mathbf{I} \cdot \mathbf{c_2} + \frac{1}{2}\Delta t^2 \cdot & \mathbf{I} \cdot \mathbf{c_3} \\ \mathbf{x_2}(t) & = & \mathbf{0} \cdot \mathbf{c_1} & + & \mathbf{I} \cdot \mathbf{c_2} + \Delta t \cdot & \mathbf{I} \cdot \mathbf{c_3} \\ \mathbf{x_3}(t) & = & \mathbf{0} \cdot \mathbf{c_1} & + & \mathbf{0} \cdot \mathbf{c_2} + & \mathbf{I} \cdot \mathbf{c_3} \end{array} \tag{5.47}$$

which is written in matrix notation as

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c_1} \\ \mathbf{c_2} \\ \mathbf{c_3} \end{bmatrix}. \tag{5.48}$$

The initial conditions summarized in the vector $\left[\mathbf{c_1}^T\ \mathbf{c_2}^T\ \mathbf{c_3}^T\right]^T$ are given by the state vector $\mathbf{x}$ at time t_0

$$\begin{array}{lcllcl} \mathbf{x_1}(t_0 = 0) & = & \mathbf{c_1} & = & \mathbf{x_0}, & \\ \mathbf{x_2}(t_0 = 0) & = & \mathbf{c_2} & = & \dot{\mathbf{x}}_0, & \text{thus} \quad \begin{bmatrix} \mathbf{c_1} \\ \mathbf{c_2} \\ \mathbf{c_3} \end{bmatrix} = \mathbf{x}(t_0). \\ \mathbf{x_3}(t_0 = 0) & = & \mathbf{c_3} & = & \ddot{\mathbf{x}}_0, & \end{array} \tag{5.49}$$

Thus the solution to (5.2) is

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \mathbf{x}(t_0) = \mathbf{\Phi}(t, t_0) \cdot \mathbf{x}(t_0). \tag{5.50}$$

For any $\Delta t = (t_i - t_{i-1})$

$$\mathbf{x}(t_i) = \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \mathbf{x}(t_{i-1}) = \mathbf{\Phi}(t_i, t_{i-1}) \cdot \mathbf{x}(t_{i-1}). \quad (5.51)$$

Alternatively the transition matrix $\mathbf{\Phi}(t, t_0)$ could have been obtained by the Laplace transformation, as shown in Section 5.3.1:

$$\begin{aligned} \mathbf{\Phi}(s) &= (s \cdot \mathbf{I} - \mathbf{F})^{-1} \\ &= \left(\begin{bmatrix} s \cdot \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & s \cdot \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & s \cdot \mathbf{I} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \right)^{-1} \\ &= \left(\begin{bmatrix} s \cdot \mathbf{I} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & s \cdot \mathbf{I} & -\mathbf{I} \\ \mathbf{0} & & s \cdot \mathbf{I} \end{bmatrix} \right)^{-1} = \begin{bmatrix} \frac{1}{s}\mathbf{I} & \frac{1}{s^2}\mathbf{I} & \frac{1}{s^3}\mathbf{I} \\ \mathbf{0} & \frac{1}{s}\mathbf{I} & \frac{1}{s^2}\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \frac{1}{s}\mathbf{I} \end{bmatrix}, \end{aligned} \quad (5.52)$$

$$\mathbf{\Phi}(t) = \mathcal{L}^{-1}\{\mathbf{\Phi}(s)\} = \begin{bmatrix} \mathbf{I} & t \cdot \mathbf{I} & \frac{1}{2}t^2\mathbf{I} \\ \mathbf{0} & \mathbf{I} & t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad (5.53)$$

$$\mathbf{\Phi}(\Delta t) = \mathbf{\Phi}(t, t_0) = \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2\mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad (5.54)$$

which is in fact equivalent to (5.50).

In order to find the inhomogenous solution of the differential equation (5.4), the matrix superposition integral is employed for $\Delta t = t - t_0$, [15],

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0) \cdot \mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau) \cdot \mathbf{G}(\tau)\mathbf{w}(\tau)d\tau + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau) \cdot \mathbf{L}(\tau)\mathbf{u}(\tau)d\tau. \quad (5.4)$$

For the differential equation in matrix notation (5.45) considering $\mathbf{u}(t) = 0$ and $\tau \in [t_0, t]$: $\mathbf{w}$ constant

$$\begin{aligned}
\mathbf{x}(t) &= \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \mathbf{x}(t_0) + \\
&\quad + \int_{t_0}^{t} \begin{bmatrix} \mathbf{I} & \tau \cdot \mathbf{I} & \frac{1}{2}\tau^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \tau \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix} d\tau \cdot \mathbf{w}(t) = \\
&= \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \mathbf{x}(t_0) + \int_{t_0}^{t} \begin{bmatrix} \frac{1}{2}\tau^2 \cdot \mathbf{I} \\ \tau \cdot \mathbf{I} \\ \mathbf{I} \end{bmatrix} d\tau \cdot \mathbf{w}(t) = \\
&= \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \mathbf{x}(t_0) + \begin{bmatrix} \frac{1}{6}\Delta t^3 \cdot \mathbf{I} \\ \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \Delta t \cdot \mathbf{I} \end{bmatrix} \cdot \mathbf{w}(t)
\end{aligned}$$

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0) \cdot \mathbf{x}(t_0) + \mathbf{\Gamma}(t, t_0) \cdot \mathbf{w}(t). \tag{5.55}$$

The system noise matrix product $\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1)$ can be obtained by using the formulation given in [15, p. 74f]:

$$\begin{aligned}
&\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) = \\
&E\Big\{ \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} \mathbf{\Phi}(t_i, \tau) \cdot \mathbf{G}(\tau)\mathbf{w}(\tau)\mathbf{w}^T(s)\mathbf{G}^T(s) \cdot \mathbf{\Phi}^T(t_i, s)\, d\tau\, ds \Big\},
\end{aligned} \tag{5.56}$$

taking the expectation operator inside the integral (proof can be found in [15]), gives

$$\begin{aligned}
&\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) = \\
&\int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} \mathbf{\Phi}(t_i, \tau) \cdot \mathbf{G}(\tau) E\Big\{ \mathbf{w}(\tau)\mathbf{w}^T(s) \Big\} \mathbf{G}^T(s) \cdot \mathbf{\Phi}^T(t_i, s)\, d\tau\, ds,
\end{aligned} \tag{5.57}$$

where the covariance of the white noise $\mathbf{w}(t)$ is defined as

$$E\{\mathbf{w}(\tau)\mathbf{w}^T(s)\} = \mathbf{Q}(\tau)\delta(\tau - s), \tag{5.58}$$

operator δ being the Dirac delta function. The properties of the Dirac delta function are used to integrate over s, giving

$$\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) = \int_{t_{i-1}}^{t_i} \mathbf{\Phi}(t_i,\tau)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}^T(\tau)\mathbf{\Phi}^T(t_i,\tau)d\tau. \tag{5.59}$$

Using the definitions for $\mathbf{\Phi}(t,\tau)$ and $\mathbf{G}(\tau)$ as in (5.55) the system noise matrix product can be obtained as

$$\begin{aligned}
&\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) = \\
&= \int_{t_{i-1}}^{t_i} \begin{bmatrix} \mathbf{I} & \tau\cdot\mathbf{I} & \frac{1}{2}\tau^2\cdot\mathbf{I} \\ \mathbf{0} & \mathbf{I} & \tau\cdot\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix} \cdot \mathbf{Q}(\tau) \cdot \\
&\qquad \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tau\cdot\mathbf{I} & \mathbf{I} & \mathbf{0} \\ \frac{1}{2}\tau^2\cdot\mathbf{I} & \tau\cdot\mathbf{I} & \mathbf{I} \end{bmatrix} d\tau \\
&= \int_{t_{i-1}}^{t_i} \begin{bmatrix} \frac{1}{2}\tau^2\cdot\mathbf{I} \\ \tau\cdot\mathbf{I} \\ \mathbf{I} \end{bmatrix} \cdot \mathbf{Q}(\tau) \cdot \begin{bmatrix} \frac{1}{2}\tau^2\cdot\mathbf{I} & \tau\cdot\mathbf{I} & \mathbf{I} \end{bmatrix} d\tau \\
&= \int_{t_{i-1}}^{t_i} \begin{bmatrix} \frac{1}{4}\tau^4\cdot\mathbf{Q}(\tau) & \frac{1}{2}\tau^3\cdot\mathbf{Q}(\tau) & \frac{1}{2}\tau^2\cdot\mathbf{Q}(\tau) \\ \frac{1}{2}\tau^3\cdot\mathbf{Q}(\tau) & \tau^2\cdot\mathbf{Q}(\tau) & \tau\cdot\mathbf{Q}(\tau) \\ \frac{1}{2}\tau^2\cdot\mathbf{Q}(\tau) & \tau\cdot\mathbf{Q}(\tau) & \mathbf{Q}(\tau) \end{bmatrix} d\tau.
\end{aligned} \tag{5.60}$$

As Gelb points out in [15, p. 75], it is important to differentiate between the two matrices $\mathbf{Q}(t)$ and $\mathbf{Q}_{i-1}$: $\mathbf{Q}(t)$ is the spectral density matrix whereas $\mathbf{Q}_{i-1}$ is the covariance matrix. A spectral density matrix may be converted to a covariance matrix through multiplication by the Dirac delta function, $\delta(t-\tau)$. Since the delta function has units of *[1/times]*, it follows that the units of the two original matrices are different.

In practical applications the term $\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1)$ can be obtained directly by error propagation, given $\mathbf{\Gamma}(i-1)$ as obtained in

(5.45) and a given definition for $\mathbf{Q}_{i-1}$:

$$\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) = \begin{bmatrix} \frac{1}{36}\Delta t^6 \cdot \mathbf{Q}_{i-1} & \frac{1}{12}\Delta t^5 \cdot \mathbf{Q}_{i-1} & \frac{1}{6}\Delta t^4 \cdot \mathbf{Q}_{i-1} \\ \frac{1}{12}\Delta t^5 \cdot \mathbf{Q}_{i-1} & \frac{1}{4}\Delta t^4 \cdot \mathbf{Q}_{i-1} & \frac{1}{2}\Delta t^3 \cdot \mathbf{Q}_{i-1} \\ \frac{1}{6}\Delta t^4 \cdot \mathbf{Q}_{i-1} & \frac{1}{2}\Delta t^3 \cdot \mathbf{Q}_{i-1} & \Delta t^2 \cdot \mathbf{Q}_{i-1} \end{bmatrix}. \tag{5.61}$$

5.5 Adaptive Kalman Filter for a kinematic network

In the following, an adaptive or self-tuning KF for the estimation of a deforming 3D geodetic network is developed, based on the innovation analysis discussed in Section 5.3.4. Detailed formulas and derivations of the KF have been given in the previous sections. All necessary terms for the KF setup are now summarized and additional assumptions are defined.

The KF needs to adapt to deformations in the network, which are not known before and may change over time. The KF is started with very rough approximate state parameter values (i.e. 0 for the states for velocity and acceleration, see also below), as this guess represents the best a priori information available. Therefore the system description is badly known. If this situation is not appropriately reflected in the stochastic system model, the system is unable to adapt to data containing information differing from the assumed starting values, possibly resulting in divergence. If the uncertainty in the system model is increased by a sufficiently large system noise error, it is able to let the estimates adapt within its range. The system uncertainty should be large enough to let the system adapt to a new situation if deformation is expected. If there is no deformation (or more generally: unexpected change in the system), the system uncertainty (i.e. system noise error) should be small, not to degrade the state estimation.

The a priori system noise error is initialized with small values and the KF determines itself, if the system noise error has to be increased, based on the comparison between new measurement data and the system description (i.e. innovation analysis). A discrepancy between the measurement data and the system description can indicate a wrong system description.

Gross errors in the observation data also result in a discrepancy between the system description and measurements. It is in fact possible to detect and identify unexpected point movements and observation gross errors in the same way. The stochastic models are iteratively adapted until system description and measurement data agree. This stochastic stabilization method has already been briefly outlined in Section 5.3.4.

If the measurement configurations are changing in a geodetic network it is necessary to keep track of the structure of the state vector. For this, the terms of observability and disturbability, discussed in Section 5.3.5 are important as they define, if a state is observable (i.e. it is possible to derive an estimate update) and disturbable. Non-disturbable parameters cannot take part in the innovation analysis and stochastic stabilization process, as a change in the related noise uncertainty does not have any effect on them.

The Kalman gain matrix $\mathbf{K}$ is the 'core' of the KF algorithm. Its structure is analyzed in Section 5.5.5. The system noise and measurement variances are the only two terms that can be chosen to influence individual elements in the gain matrix thus providing the basis for stochastic stabilization to work.

The section is concluded with a short discussion of a simple method to decide on the correct system model description for an application, Section 5.5.6.

5.5.1 Summary of KF terms for estimation of a kinematic 3D network and assumptions for the stochastic models

The state vector is interpreted as a composite: It contains for each point in a 3D network position, velocity and acceleration coordinate components. In general it might be possible to observe all of these state variables directly (e.g. by including acceleration measurements etc.) but in traditional surveying applications, only the position coordinate part of the state vector is directly observed (see observability in Section 5.3.5). Nevertheless velocity and acceleration states can also be determined as the system is fully disturbable (see Section 5.5.4). An extended KF applies as the observation equations in survey applications are generally non-linear. The state vector contains the absolute state variables, as has been discussed in Section 5.3.3.

- Initialization of state information in zero (or pre-zero) epoch: It is assumed that no information on velocity and acceleration values is available, but as these are considered to be small they are approximated by 0. No correlations between the state variables are assumed.

$$\hat{\mathbf{x}}_{0|0} = \begin{bmatrix} \hat{\mathbf{x}}_{0|0} \\ \dot{\hat{\mathbf{x}}}_{0|0} \\ \ddot{\hat{\mathbf{x}}}_{0|0} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}_{0|0} \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{P}_{0|0} = \mathrm{diag}(\mathbf{p}_0). \tag{5.62}$$

- Time difference between epochs, not necessarily equidistant:

$$\Delta t = t_i - t_{i-1}. \tag{5.63}$$

- System equations for unforced, uniformly accelerated motion:

$$\hat{\mathbf{x}}_{i|i-1} = \mathbf{\Phi}(i, i-1)\hat{\mathbf{x}}_{i-1|i-1}, \tag{5.14}$$

$$\mathbf{P}_{i|i-1} = \mathbf{\Phi}(i,i-1)\mathbf{P}_{i-1|i-1}\mathbf{\Phi}^T(i,i-1) + \mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1). \tag{5.15}$$

- Transition matrix for polynomial approximation of motion:

$$\mathbf{\Phi}(i, i-1) = \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{5.64}$$

- System noise uncertainty (matrix product), assuming a white system noise process:

$$\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) = \begin{bmatrix} \frac{1}{36}\Delta t^6 \cdot \mathbf{Q}_{i-1} & \frac{1}{12}\Delta t^5 \cdot \mathbf{Q}_{i-1} & \frac{1}{6}\Delta t^4 \cdot \mathbf{Q}_{i-1} \\ \frac{1}{12}\Delta t^5 \cdot \mathbf{Q}_{i-1} & \frac{1}{4}\Delta t^4 \cdot \mathbf{Q}_{i-1} & \frac{1}{2}\Delta t^3 \cdot \mathbf{Q}_{i-1} \\ \frac{1}{6}\Delta t^4 \cdot \mathbf{Q}_{i-1} & \frac{1}{2}\Delta t^3 \cdot \mathbf{Q}_{i-1} & \Delta t^2 \cdot \mathbf{Q}_{i-1} \end{bmatrix}. \tag{5.61}$$

- Covariance of system noise $\mathbf{w}_{i-1}$, assumed to be uncorrelated:

$$\mathbf{Q}_{i-1} = \mathrm{diag}(\mathbf{q}_{i-1}). \tag{5.65}$$

- Update of state vector by including measurements $\mathbf{z}_i$ available in epoch i:

$$\hat{\mathbf{x}}_{i|i} = \hat{\mathbf{x}}_{i|i-1} + \mathbf{K}_i(\mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}) = \hat{\mathbf{x}}_{i|i-1} + \mathbf{K}_i\mathbf{d}_i. \quad (5.66)$$

- Updated state covariance:

$$\mathbf{P}_{i|i} = (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{P}_{i|i-1}. \quad (5.10)$$

- Gain matrix:

$$\mathbf{K}_i = \mathbf{P}_{i|i-1}\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i)^{-1}. \quad (5.19)$$

- Linearized matrix of observation equations – (Assumption: only measurements for position components):

$$\mathbf{H_i} = \left[\begin{bmatrix} \frac{\delta \mathbf{h_1}}{\delta \mathbf{x1}} & \frac{\delta \mathbf{h_1}}{\delta \mathbf{x2}} & \cdots & \frac{\delta \mathbf{h_1}}{\delta \mathbf{x}u} \\ \frac{\delta \mathbf{h_2}}{\delta \mathbf{x1}} & \frac{\delta \mathbf{h_2}}{\delta \mathbf{x2}} & & \vdots \\ \vdots & & \ddots & \\ \frac{\delta \mathbf{h_n}}{\delta \mathbf{x1}} & \cdots & & \frac{\delta \mathbf{h_n}}{\delta \mathbf{x}m} \end{bmatrix} \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & 0 & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & 0 & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}\right]$$

$$\mathbf{H_i} = \begin{bmatrix} \mathbf{H}_{i\,pos} & \mathbf{H}_{i\,vel} & \mathbf{H}_{i\,acc} \end{bmatrix}. \quad (5.67)$$

- Covariance of uncorrelated measurements:

$$\mathbf{R}_i = \mathrm{diag}(\mathbf{r}_i). \quad (5.68)$$

- Vector of innovations and its covariance:

$$\mathbf{d}_i = \mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1}, \quad (5.21)$$

$$\mathbf{D}_i = \mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i. \quad (5.22)$$

5.5.2 Innovation analysis for model error detection and identification

As already discussed in Section 5.3.4 and at the beginning of this section, the KF can be set up to adapt to new information, previously unforeseen in the system description, by adaptive methods. A Gauss-Markov (GM) model interpretation of the KF equations additionally enables not only to search for misspecifications in the stochastic system noise model but also in the stochastic measurement model. It has been briefly argued in Section 5.3.4 that in the application of a geodetic network the correlation methods applied to innovations are not suited to give reasonable results. This is due to the characteristics of measurement data of networks in changing configurations giving very small and incomplete sample sizes resulting in poor statistics. As a consequence the less demanding but also less conclusive method of stochastic stabilization is chosen to adapt the filter for correct identification of the stochastic models.

It has to be ensured that the system is completely disturbable. The term of disturbability has been defined in Section 5.3.5 and in the following Section 5.5.4 it is derived for the example at hand, showing that this condition is satisfied after a certain initialization lag.

Gauss-Markov (GM) model of the KF [19] and [63] interpret the KF as an extended GM model. The predicted state vector $\hat{\mathbf{x}}_{i|i-1}$ is considered to be part of the observation data. Together with the actual observations $\mathbf{z}_i$ they are used to estimate the updated state vector $\hat{\mathbf{x}}_{i|i}$. The GM model in the form $\mathbf{l} + \mathbf{v} = \mathbf{A}\hat{\mathbf{x}}$ transforms to

$$\begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{z}_i \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{\mathbf{x},i} \\ \mathbf{v}_{\mathbf{z},i} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_i \end{bmatrix} \hat{\mathbf{x}}_{i|i}. \tag{5.69}$$

The residuals $\mathbf{v}_{\mathbf{x},i}$ and $\mathbf{v}_{\mathbf{z},i}$ can be derived as

$$\begin{aligned} \begin{bmatrix} \mathbf{v}_{\mathbf{x},i} \\ \mathbf{v}_{\mathbf{z},i} \end{bmatrix} &= \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_i \end{bmatrix} \hat{\mathbf{x}}_{i|i} - \begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{z}_i \end{bmatrix} \\ &= \begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} + \mathbf{K}_i\mathbf{d}_i \\ \mathbf{H}_i\hat{\mathbf{x}}_{i|i-1} + \mathbf{H}_i\mathbf{K}_i\mathbf{d}_i \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{z}_i \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{K}_i\mathbf{d}_i \\ \mathbf{H}_i\mathbf{K}_i\mathbf{d}_i - \mathbf{d}_i \end{bmatrix} \\ \begin{bmatrix} \mathbf{v}_{\mathbf{x},i} \\ \mathbf{v}_{\mathbf{z},i} \end{bmatrix} &= \begin{bmatrix} \mathbf{K}_i \\ (\mathbf{H}_i\mathbf{K}_i - \mathbf{I}) \end{bmatrix} \mathbf{d}_i. \end{aligned} \tag{5.70}$$

The predicted state vector $\hat{\mathbf{x}}_{i|i-1}$ and the observation data $\mathbf{z}_i$ for time i are uncorrelated, the joint covariance matrix is

$$\mathbf{R}_{\mathbf{xz},i} = \begin{bmatrix} \mathbf{P}_{i|i-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_i \end{bmatrix}. \tag{5.71}$$

The covariance matrix of the residuals follows by error propagation of (5.69) as

$$\mathbf{V}_{GM,i} = \begin{bmatrix} \mathbf{P}_{i|i-1}\mathbf{H}_i^T\mathbf{K}_i^T & \mathbf{P}_{i|i-1}\mathbf{H}_i^T\left(\mathbf{K}_i^T\mathbf{H}_i^T - \mathbf{I}\right) \\ \left(\mathbf{H}_i\mathbf{K}_i - \mathbf{I}\right)\mathbf{H}_i\mathbf{P}_{i|i-1} & \left(\mathbf{H}_i\mathbf{K}_i - \mathbf{I}\right)\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i \end{bmatrix}. \tag{5.72}$$

Equivalently to (5.70) the residuals can be derived as, [19]:

$$\begin{aligned} \mathbf{v}_{GM,i} &= -\mathbf{H}_{GM,i} \cdot \mathbf{z}_{GM,i} \\ \begin{bmatrix} \mathbf{v}_{\mathbf{x},i} \\ \mathbf{v}_{\mathbf{z},i} \end{bmatrix} &= \begin{bmatrix} -\mathbf{K}_i\mathbf{H}_i & \mathbf{K}_i \\ \mathbf{R}_i\mathbf{D}_i^{-1}\mathbf{H}_i & -\mathbf{R}_i\mathbf{D}_i^{-1} \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{x}}_{i|i-1} \\ \mathbf{z}_i \end{bmatrix}. \end{aligned} \tag{5.73}$$

The global model test also referred to as *local overall model* test, [63]. This test decides whether a significant difference between the system model description and the measurement data exists. It makes use of the innovation property as defined in (5.33) and (5.34). Errors can be located in the predicted state vector and the measurements, and need to be identified in a following localization step.

The innovation is considered to be a gaussian distributed variable, [35, p. 228]:

$$H_0 \; : \; \mathbf{d}_i \sim N(\mathbf{0}, \mathbf{D}_i), \tag{5.74}$$

which represents the null hypothesis of the global model test for the KF process. The quadratic form $\mathbf{d}_i^T\mathbf{D}_i^{-1}\mathbf{d}_i$ is χ^2-distributed. A test term can be defined

$$\Omega^2_{\mathbf{d}_i} = \mathbf{d}_i^T\mathbf{D}_i^{-1}\mathbf{d}_i \sim \chi^2_{1-\alpha,f_i}, \tag{5.75}$$

with the probability relationship

$$P\{\Omega^2_{\mathbf{d}_i} \leq \chi^2_{1-\alpha,f_i} | H_0\} = 1 - \alpha, \tag{5.76}$$

where f_i is the degree of freedom determining $\mathbf{d}_i$. If the null hypothesis H_0 is rejected at a significance level of $(1-\alpha)$, the observations carry information different from the system description.

A rejection of the null hypothesis indicates either an incorrect functional or stochastic model. The above test formulation is similar to the global congruency model test in classical deformation analysis which determines if significant coordinate differences between two observation epochs (i.e. deformations) have occurred, as already discussed in Chapter 4.2.

Localization To search for the misspecifications in the predicted state vector or measurement data indicated by the innovation analysis in the global model test, the individual terms are analyzed. The corresponding stochastic information is adapted iteratively until the global test of innovations is passed. The derivation of the following test statistics is based on the approach by Pelzer in [49] which I extended here for the GM model interpretation as summarized in (5.73).

If the null hypothesis H_0 defined in (5.74) is valid, it can be reformulated as

$$H_0 \; : \; \mathbf{v}_{GM,i} \sim N(\mathbf{0}, \mathbf{V}_{GM,i}). \tag{5.77}$$

The test statistic similar to (5.75) can be written as

$$\Omega^2_{\mathbf{v}_{GM,i}} = \mathbf{v}^T_{GM,i}\mathbf{V}^{-1}_{GM,i}\mathbf{v}_{GM,i} \sim \chi^2_{1-\alpha,h_i}, \tag{5.78}$$

with $h_i = \mathrm{rank}(\mathbf{V}_{GM,i})$, or more generally

$$\Omega^2_{\mathbf{v}_{GM,i}} = \mathbf{v}^T_{GM,i}\mathbf{V}^{+}_{GM,i}\mathbf{v}_{GM,i} \sim \chi^2_{1-\alpha,h_i}. \tag{5.79}$$

To analyze individual components of $\mathbf{v}_{GM,i}$ for a significant outlier, test statistics have to be defined. The components of the predicted state vector residuals $\mathbf{v}_{x,i}$ can be grouped as components of 3D points they refer to, e.g. entries for position, velocity and acceleration states for a point j.

Pelzer suggested in [49] that state parameters describing a single point are very closely related and should be investigated together concerning possible point movements. In the case of geodetic network measurements a decoupling between horizontal and vertical observation information can be seen (e.g. measurements by hydrostatic levelling system (HLS) and offset measurements (WPS), see Chapter 3.3). In this

case it is indicated to form individual groups for horizontal and vertical components of one point and to analyze these two groups separately. This also applies if point deformations are expected to be dominantly either in horizontal or vertical direction. As the grouping of horizontal and vertical components only represents a further specialization of the case of the 3D point j, this notation is adhered to in the following, allowing for a more compact formulation.

Omitting the time index i the components for point j ($j \leq$ number of points) are:

$$\mathbf{v}_{GM,j} = \mathbf{E}_{[j]}\mathbf{v}_{GM} = \begin{bmatrix} \mathbf{v}_{x,j\,pos} \\ \mathbf{v}_{x,j\,vel} \\ \mathbf{v}_{x,j\,acc} \end{bmatrix}, \tag{5.80}$$

where $\mathbf{E}_{[j]}$ is a matrix of unity row vectors, with values 1 corresponding to indices in $\mathbf{v}_{GM,j}$ referring to entries for point j. If m elements refer to point j, $\mathbf{E}_{[j]}$ has m rows.

The covariance matrix $\mathbf{V}_{GM,j}$ for point j is obtained correspondingly:

$$\mathbf{V}_{GM,j} = \mathbf{E}_{[j]}\mathbf{V}_{GM}\mathbf{E}_{[j]}^T = \begin{bmatrix} \mathbf{V}_{GM,j\,pos|j\,pos} & \mathbf{V}_{GM,j\,pos|j\,vel} & \mathbf{V}_{GM,j\,pos|j\,acc} \\ \mathbf{V}_{GM,j\,vel|j\,pos} & \mathbf{V}_{GM,j\,vel|j\,vel} & \mathbf{V}_{GM,j\,vel|j\,acc} \\ \mathbf{V}_{GM,j\,acc|j\,pos} & \mathbf{V}_{GM,j\,acc|j\,vel} & \mathbf{V}_{GM,j\,acc|j\,acc} \end{bmatrix}. \tag{5.81}$$

Measurement residuals $\mathbf{v}_{z,i}$ are analyzed individually. Thus in the following they are analyzed in 'groups of 1' to allow for a unified formulation. For ($j >$ number of points):

$$\mathbf{v}_{GM,j} = \mathbf{E}_{[j]}\mathbf{v}_{GM} = \begin{bmatrix} \mathbf{v}_{z,j} \end{bmatrix}, \quad \mathbf{V}_{GM,j} = \mathbf{E}_{[j]}\mathbf{V}_{GM}\mathbf{E}_{[j]}^T. \tag{5.82}$$

In this case the matrix $\mathbf{E}_{[j]}$ is actually a single row vector.

For each point j a test statistic similar to (5.78) is formulated:

$$\Omega^2_{\mathbf{v}_{GM,j}} = \mathbf{v}_{GM,j}^T\mathbf{V}_{GM,j}^{-1}\mathbf{v}_{GM,j} \sim \chi^2_{1-\bar{\alpha},h_j}, \tag{5.83}$$

with $h_j = \text{rank}(\mathbf{V}_{GM,j})$ and the probability relationship

$$P\{\Omega^2_{\mathbf{v}_{GM,j}} \leq \chi^2_{1-\bar{\alpha},h_j} | H_0\} = 1 - \bar{\alpha}, \tag{5.84}$$

where $1 - \bar{\alpha}$ is the adapted level of test significance defined later in this section.

Strict testing In (5.81) – (5.83) any correlations between different points are neglected. This can bias the test decision. The correct way described in the following is referred to as *strict* or rigorous testing, [19].

Splitting up the vector $\mathbf{v}_{GM}$ (time index i again suppressed) in one part $\mathbf{v}_{GM,1}$ of unsuspicious points' entries and a second part $\mathbf{v}_{GM,2}$ corresponding to the point suspicious of deformation:

$$\mathbf{v}_{GM} = \begin{bmatrix} \mathbf{v}_{GM,1} \\ \mathbf{v}_{GM,2} \end{bmatrix}. \tag{5.85}$$

The inverse of the covariance matrix $\mathbf{V}_{GM}$ is split up accordingly:

$$\mathbf{V}_{GM}^{-1} = \begin{bmatrix} \mathbf{iV}_{GM,11} & \mathbf{iV}_{GM,12} \\ \mathbf{iV}_{GM,21} & \mathbf{iV}_{GM,22} \end{bmatrix}. \tag{5.86}$$

The correct test statistic equivalent to (5.83) is formulated in the following way:

$$\mathbf{v}_{GM}^T \mathbf{V}_{GM}^{-1} \mathbf{v}_{GM} = \mathbf{v}_{GM}^T \mathbf{F}_1^T \mathbf{F}_2 \mathbf{V}_{GM}^{-1} \mathbf{F}_2^T \mathbf{F}_1 \mathbf{v}_{GM} = \overline{\mathbf{v}}_{GM}^T \overline{\mathbf{V}}_{GM}^{-1} \overline{\mathbf{v}}_{GM}, \tag{5.87}$$

where the transformation matrices $\mathbf{F}_1$ and $\mathbf{F}_2$ fulfill the condition:

$$\mathbf{F}_1^T \mathbf{F}_2 = \mathbf{F}_2^T \mathbf{F}_1 = \mathbf{I}. \tag{5.88}$$

Thus the transformed terms are obtained as:

$$\overline{\mathbf{v}}_{GM} = \mathbf{F}_1 \mathbf{v}_{GM} \tag{5.89}$$

$$\overline{\mathbf{V}}_{GM}^{-1} = \mathbf{F}_2 \mathbf{V}_{GM}^{-1} \mathbf{F}_2^T. \tag{5.90}$$

Defining $\mathbf{F}_1$ as, [19, p. 139]:

$$\mathbf{F}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{iV}_{GM,22}^{-1} \mathbf{iV}_{GM,21} & \mathbf{I} \end{bmatrix}, \tag{5.91}$$

and considering the condition (5.88) gives for $\mathbf{F}_2$

$$\mathbf{F}_2 = \begin{bmatrix} \mathbf{I} & -\mathbf{iV}_{GM,12} \mathbf{iV}_{GM,22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{5.92}$$

This gives the transformed terms as:

$$\begin{bmatrix} \overline{\mathbf{v}}_{GM,1} \\ \overline{\mathbf{v}}_{GM,2} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_{GM,1} \\ \mathbf{iV}_{GM,22}^{-1} \mathbf{iV}_{GM,21} \mathbf{v}_{GM,1} + \mathbf{v}_{GM,2} \end{bmatrix} \tag{5.93}$$

$$\begin{aligned}\overline{\mathbf{V}}_{GM}^{-1} &= \begin{bmatrix} \mathbf{iV}_{GM,11} - \mathbf{iV}_{GM,12}\mathbf{iV}_{GM,22}^{-1}\mathbf{iV}_{GM,21} & \mathbf{0} \\ \mathbf{0} & \mathbf{iV}_{GM,22} \end{bmatrix} \\ &= \begin{bmatrix} \overline{\mathbf{iV}}_{GM,11} & \mathbf{0} \\ \mathbf{0} & \mathbf{iV}_{GM,22} \end{bmatrix}. \end{aligned} \tag{5.94}$$

The original test statistic as defined in (5.79) becomes

$$\begin{aligned} \mathbf{v}_{GM}^T\mathbf{V}_{GM}^{-1}\mathbf{v}_{GM} &= \mathbf{v}_{GM,1}^T\overline{\mathbf{iV}}_{GM,11}\mathbf{v}_{GM,1} + \overline{\mathbf{v}}_{GM,2}^T\mathbf{iV}_{GM,22}\overline{\mathbf{v}}_{GM,2} \\ \Omega^2 &= \Omega_1^2 + \Omega_R^2 \end{aligned} \tag{5.95}$$

where Ω_R is the additional term due to the suspicious part of $\mathbf{v}_{GM}$, i.e. $\mathbf{v}_{GM,2}$.

With

$$\overline{\mathbf{v}}_{GM,2} = \mathbf{iV}_{GM,22}^{-1}\mathbf{iV}_{GM,21}\mathbf{v}_{GM,1} + \mathbf{v}_{GM,2} \tag{5.96}$$

the rigorous test statistic becomes

$$\Omega_R^2 = \overline{\mathbf{v}}_{GM,2}^T\mathbf{iV}_{GM,22}\overline{\mathbf{v}}_{GM,2} \sim \chi^2_{1-\bar{\alpha},h_2}, \tag{5.97}$$

which represents the test for the suspicious part of $\mathbf{v}_{GM}$ or as before in (5.83) formulated for each point j:

$$\Omega^2_{\overline{\mathbf{v}}_{GM,j}} = \overline{\mathbf{v}}_{GM,j}^T\mathbf{iV}_{GM,jj}\overline{\mathbf{v}}_{GM,j} \sim \chi^2_{1-\bar{\alpha},h_j}, \tag{5.98}$$

with

$$\overline{\mathbf{v}}_{GM,j} = \mathbf{iV}_{GM,jj}^{-1}\mathbf{iV}_{GM,jk}\mathbf{v}_{GM,k} + \mathbf{v}_{GM,j} \tag{5.99}$$

$$\mathbf{v}_{GM} = \begin{bmatrix} \mathbf{v}_{GM,k} \\ \mathbf{v}_{GM,j} \end{bmatrix} \begin{matrix} \text{unsuspicious part} \\ \text{suspicious entries} \end{matrix} \tag{5.100}$$

$$\mathbf{V}_{GM}^{-1} = \begin{bmatrix} \mathbf{iV}_{GM,kk} & \mathbf{iV}_{GM,kj} \\ \mathbf{iV}_{GM,jk} & \mathbf{iV}_{GM,jj} \end{bmatrix}. \tag{5.101}$$

Adapted level of test significance In (5.83) to (5.98) the term $1-\bar{\alpha}$ describes the adapted level of the test significance for a single test within a series of n repeated tests. This adaptation is advisable if the results of the individual tests should correspond to the result of the global test (5.79) as these n repeated tests are strictly speaking not independent.

As an approximation for the error probability $\bar{\alpha}$ serves, [44, p. 281]:

$$\bar{\alpha} \approx \frac{\alpha}{n}, \tag{5.102}$$

where n represents the number of single tests, i.e. in the case discussed above it corresponds to the number of points. Niemeier points out in [44, p. 281] that if the network becomes larger, and n increases, $\bar{\alpha}$ becomes smaller and thus the threshold for the statistical test increases. He suggested to use the redundancy R (defined in Chapter 4.1) for a more reliable estimate of $\bar{\alpha}$,

$$\bar{\alpha} \approx \frac{\alpha}{R}. \tag{5.103}$$

$\bar{\alpha}$ is smaller than α, thus the level of significance $1-\bar{\alpha}$ is increased. This has the effect that an error or outlier in the test is not so easily suspected as with α.

Heunecke determines in [19] the redundancy R in the KF in Gauss-Markov formulation (5.73) as the total system redundancy:

$$R_{sys} = \text{trace}(\mathbf{H}_{\mathbf{GM,v}}) = \sum r_k = n_{\mathbf{z},i}, \tag{5.104}$$

where $n_{\mathbf{z},i}$ refers to the number of measurements $\mathbf{z}_i$ in epoch i. Redundancy contributions for individual observations can be obtained using the form

$$r_k = \eta_{[k]}^T \mathbf{H}_{\mathbf{GM,v}} \eta_{[k]}; \qquad k = 1, \ldots, n_i, \tag{5.105}$$

where where $\eta_{[k]}$ is a unity vector with zeros for all elements but the element k, which is 1 and n_i refers to the number of all 'observations', i.e. measurements and predicted state variables. Summarized for the two 'observation groups' $\hat{\mathbf{x}}_{i|i-1}$ and $\mathbf{z}_i$ the redundancy numbers for the predicted state vector and for the measurement data can be given

$$R_{\hat{\mathbf{x}}_{i|i-1}} = \sum_{k=1}^{n_{x,i}} \eta_{[k]}^T \mathbf{K}_i \mathbf{H}_i \eta_{[k]} = \sum_{k=1}^{n_{x,i}} r_{\hat{\mathbf{x}}_{i|i-1},k} \tag{5.106}$$

$$R_{\mathbf{z}_i} = \sum_{k=1}^{n_{z,i}} \eta_{[k]}^T \mathbf{R}_i \mathbf{D}_i^{-1} \eta_{[k]} = \sum_{k=1}^{n_{z,i}} r_{\mathbf{z}_i,k}. \tag{5.107}$$

The individual redundancy numbers $r_{\hat{\mathbf{x}}_{i|i-1},k}$ and $r_{\mathbf{z}_i,k}$ describe the contribution of the individual 'observation' to the overall redundancy of the estimation process. These individual redundancy numbers (summarized in the following expression as r_k) can be used to adapt the test significance level to make the statistical test more sensitive in case of poorly controlled observations or predicted states.

$$\bar{\alpha} \approx \alpha \cdot (1 - r_k). \tag{5.108}$$

The resulting adapted level of significance of the test $1 - \bar{\alpha}$ does not necessarily correspond to the global model test.

If the redundancy number of the corresponding predicted state or measurement is small, $\bar{\alpha}$ is close to α and generally larger than $\bar{\alpha} \approx \frac{\alpha}{R}$. Compared to the latter an outlier is more easily suspected, at the cost of a reduced significance of the test. As errors in poorly controlled observations are often projected also on other parameters, and become thus less easily detectable, it is reasonable to lower the threshold for the corresponding test statistic in these cases, thereby increasing the type I error probability.

5.5.3 Stochastic stabilization

In the previous section the test statistic (5.83) or (5.98) has been formulated to analyze whether predicted state variables and observations are the reason for the discrepancy between the system and the measurement data. A method similar to gross error detection is applied to decide which entries differ significantly from the expected values. Gross error or outlier detection is in general applied to the observation data, as discussed in Chapter 4.1, but with the interpretation of the GM model in (5.69) it can be applied also to the extended observation vector $\mathbf{z}_{GM,i}$. In the following the term *outlier* is used for a deformation in the state vector or a gross error in the observations. In case of a single assumed outlier in the 'observation' material, the method of *data snooping* introduced by Baarda, [4] is a widely used technique in geodetic applications, which defines for each observation a test statistic

of standardized residuals, as defined in Chapter 4.1. The observation with the largest value for the test statistic is rejected. This works well if only a single outlier is present, but as this assumption is usually not true, the process is iterated: in each iteration step the observation with the largest test statistic value is rejected and the adjustment repeated until no significant model mismatch is found. But Wieser stresses in [62] that in the case of more than one outlier the detection becomes difficult as the test statistics based on the standardized residuals are only to be understood as approximations. A more robust method of detecting outlying observations is obtained by an iterative re-weighted least squares estimator, [62], similar to the *Danish method*, [28]. Instead of rejecting a single observation based on its large test statistic, all observations with a significant test statistic are re-weighted (down-weighted), based on the ratio by which the test statistic exceeds the respective test threshold. Thus an observation with a large residual resulting in a large test statistic gets small weights and has less influence on the next adjustment run. An outlying observation does get down-weighted iteratively to zero and does not have any influence on the estimation process.

To achieve stochastic stabilization of the KF, I choose a similar approach as in [62]: For each group j the following ratio is derived

$$T_j = \frac{\Omega^2_{\overline{\mathbf{v}}_{GM,j}}}{\chi^2_{1-\overline{\alpha},h_j}}. \tag{5.109}$$

The stochastic model definitions are modified based on this test term. In case of the measurements, the corresponding error information $\mathbf{R}_i$ can be directly modified, assuming uncorrelated observations $\mathbf{R}_i = \mathrm{diag}(\mathbf{r}_i)$. In the following formulations the time index i is added again:

$$\begin{aligned} T_{i,j} &> 1 \qquad \mathbf{r}_i = \mathbf{r}_i \cdot \exp T_{i,j} \\ T_{i,j} &\leq 1 \qquad \mathbf{r}_i = \mathbf{r}_i. \end{aligned} \tag{5.110}$$

In case of the predicted state vector which is analyzed for outliers in the system model, the corresponding entries in the system noise error information has to be modified in order to achieve the necessary stabilization. $\mathbf{Q}_{i-1}$ is changed, assuming uncorrelated system noise $\mathbf{Q}_{i-1} = \mathrm{diag}(\mathbf{q}_{i-1})$.

$$\begin{aligned} T_{i,j} &> 1 \qquad \mathbf{q}_{i-1} = \mathbf{q}_{i-1} \cdot \exp T_{i,j} \\ T_{i,j} &\leq 1 \qquad \mathbf{q}_{i-1} = \mathbf{q}_{i-1}. \end{aligned} \tag{5.111}$$

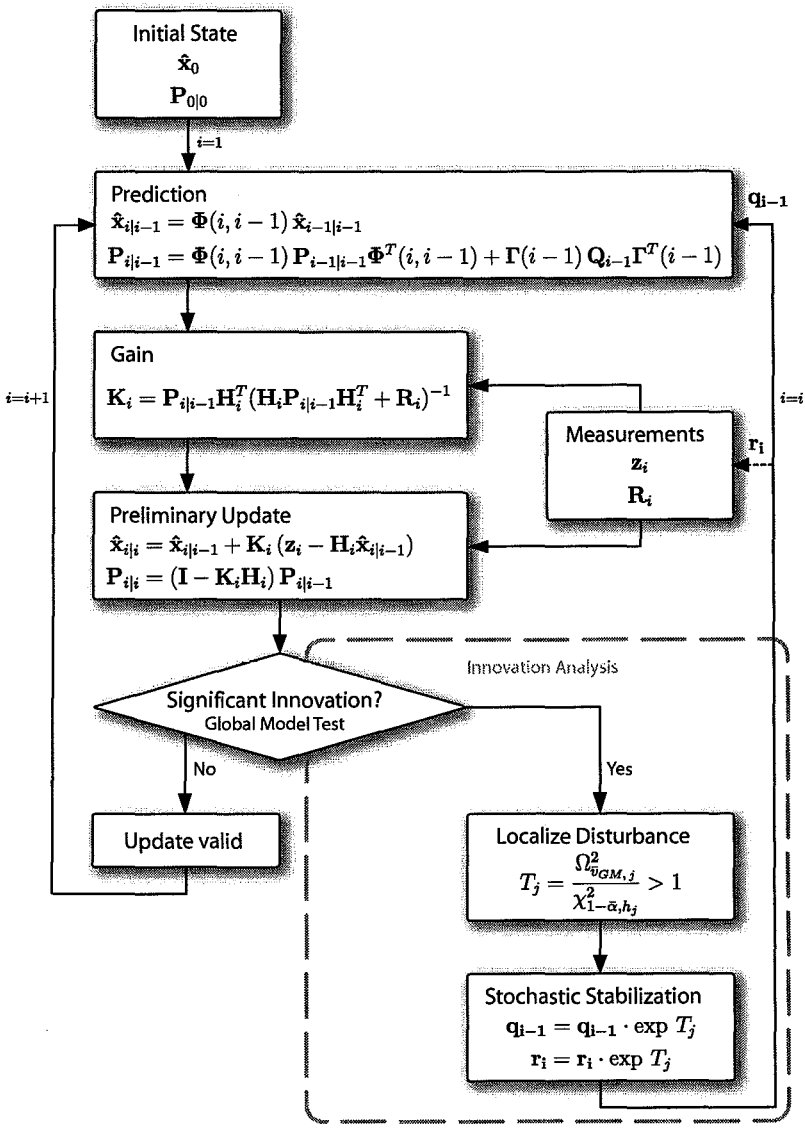


Figure 5.3: *The adaptive Kalman Filter.*

The estimation process is repeated. If the global model test (5.76) does still indicate a significant system disturbance, the stochastic stabilization is repeated. This iterative process terminates, once no significant difference between the measurement data and the system can be identified.

The flow of processes in the adaptive KF described here is summarized in Figure 5.3 and comparison with Figure 5.2 shows the differences with the standard KF.

5.5.4 KF characteristic terms

As discussed before in Section 5.3.5 and at the beginning of this section, the terms of observability and disturbability contain important information about the system's characteristics. These two terms are now derived for the example of the kinematic KF setup for a 3D geodetic network of classical survey measurements.

Observability From (5.39), (5.64) and (5.67) and by deriving

$$\mathbf{\Phi}^T(i,0)\mathbf{H}_i^T = \begin{bmatrix} \mathbf{H}_{i\,pos}^T & & \\ (t_i - t_0)\mathbf{H}_{i\,pos}^T & +\mathbf{H}_{i\,vel}^T & \\ \frac{1}{2}(t_i - t_0)^2\mathbf{H}_{i\,pos}^T & +(t_i - t_0)\mathbf{H}_{i\,vel}^T & +\mathbf{H}_{i\,acc}^T \end{bmatrix} \tag{5.112}$$

the observability matrix $\mathbf{\Xi}$ follows as

$$\mathbf{\Xi} = \left[\begin{array}{c|c|c} \mathbf{H}_{0\,pos}^T & \mathbf{H}_{1\,pos}^T & |\cdots| \\ \mathbf{H}_{0\,vel}^T & (t_1 - t_0)\mathbf{H}_{1\,pos}^T + \mathbf{H}_{1\,vel}^T & |\cdots| \\ \mathbf{H}_{0\,acc}^T & \frac{1}{2}(t_1 - t_0)^2\mathbf{H}_{1\,pos}^T + (t_1 - t_0)\mathbf{H}_{1\,vel}^T + \mathbf{H}_{1\,acc}^T & |\cdots| \end{array}\right.$$

$$\left.\begin{array}{c} \mathbf{H}_{n\,pos}^T \\ (t_n - t_0)\mathbf{H}_{n\,pos}^T + \mathbf{H}_{n\,vel}^T \\ \frac{1}{2}(t_n - t_0)^2\mathbf{H}_{n\,pos}^T + (t_n - t_0)\mathbf{H}_{n\,vel}^T + \mathbf{H}_{n\,acc}^T \end{array}\right]. \tag{5.113}$$

This expression can be simplified as in (5.67) (because of the assumption $\mathbf{H}_{i\,vel}^T = \mathbf{H}_{i\,acc}^T = \mathbf{0}$)

$$\boldsymbol{\Xi} = \begin{bmatrix} \mathbf{H}_{0\,pos}^T & | & \mathbf{H}_{1\,pos}^T & | & \cdots & | & \mathbf{H}_{n\,pos}^T \\ \mathbf{0} & | & (t_1 - t_0)\mathbf{H}_{1\,pos}^T & | & \cdots & | & (t_n - t_0)\mathbf{H}_{n\,pos}^T \\ \mathbf{0} & | & \frac{1}{2}(t_1 - t_0)^2\mathbf{H}_{1\,pos}^T & | & \cdots & | & \frac{1}{2}(t_n - t_0)^2\mathbf{H}_{n\,pos}^T \end{bmatrix}. \tag{5.114}$$

As expected the observability is dependent on the measurement matrices $\mathbf{H}_i$. In the first epoch, $\text{rank}(\boldsymbol{\Xi}) = \text{rank}(\mathbf{H}_0) = \text{rank}(\mathbf{H}_{0\,pos}) = u - d$. Thus in the first epoch only position states are observable. In the next epoch also entries for velocity and acceleration components are created by the transition matrix, as in (5.112). But $\text{rank}(\boldsymbol{\Phi}^T(1,0) \cdot \mathbf{H}_1^T)$ is still the same as before in the first epoch, as the different blocks for the position, velocity and acceleration components are linearly dependent on each other. Added to the $\boldsymbol{\Xi}$ matrix from epoch one, however, the overall rank of the matrix changes, as the block of rows for position components and the block of velocity and acceleration components are now independent, if $(t_1 - t_0) \neq 0$. So, in the second epoch, the observability is in principle doubled (assuming that all states are observed in both epochs). Now also velocity states get update values $\neq 0$. This is repeated in the third epoch for the acceleration components to become observable and full observability can be reached, also if the individual measurements refer only to the state position components.

Disturbability From (5.41), (5.55) and (5.64) and by deriving

$$\boldsymbol{\Phi}(i,1)\boldsymbol{\Gamma}(1) = \begin{bmatrix} (\frac{1}{6}(t_1 - t_0)^3 + \frac{1}{2}(t_1 - t_0)^2(t_i - t_1) + \frac{1}{2}(t_1 - t_0)(t_i - t_1)^2) \cdot \mathbf{I} \\ (\frac{1}{2}(t_1 - t_0)^2 + (t_1 - t_0)(t_i - t_1)) \cdot \mathbf{I} \\ (t_1 - t_0) \cdot \mathbf{I} \end{bmatrix} \tag{5.115}$$

the disturbability matrix $\mathbf{\Psi}$ can be obtained as

$$\mathbf{\Psi} = \begin{bmatrix} \frac{1}{6}(t_1-t_0)^3 \cdot \mathbf{I} & | \\ \frac{1}{2}(t_1-t_0)^2 \cdot \mathbf{I} & | \\ (t_1-t_0) \cdot \mathbf{I} & | \end{bmatrix} \begin{array}{|c|} (\frac{1}{6}(t_1-t_0)^3 + \frac{1}{2}(t_1-t_0)^2(t_2-t_1) + \frac{1}{2}(t_1-t_0)(t_2-t_1)^2) \cdot \mathbf{I} \\ (\frac{1}{2}(t_1-t_0)^2 + (t_1-t_0)(t_2-t_1)) \cdot \mathbf{I} \\ (t_1-t_0) \cdot \mathbf{I} \end{array} \begin{array}{|c|} \cdots \\ \cdots \\ \cdots \end{array} \begin{array}{|c|} (\frac{1}{6}(t_1-t_0)^3 + \frac{1}{2}(t_1-t_0)^2(t_n-t_1) + \frac{1}{2}(t_1-t_0)(t_n-t_1)^2) \cdot \mathbf{I} \\ (\frac{1}{2}(t_1-t_0)^2 + (t_1-t_0)(t_n-t_1)) \cdot \mathbf{I} \\ (t_1-t_0) \cdot \mathbf{I} \end{array} \Bigg]. \tag{5.116}$$

Disturbability of a state is delayed by one epoch compared to the state's observability. The rank of the whole matrix $\mathbf{\Psi}$ is governed largely by the rank of $\mathbf{\Gamma}$ and its propagation in time by the transition matrix, very similar to the situation of observability. Each single matrix $\mathbf{\Phi}(i,1) \cdot \mathbf{\Gamma}(1)$ has a rank of one third of its dimension, but concatenated to form the matrix $\mathbf{\Psi}$ the rank increases with each added block, if $(t_i - t_0) \neq 0$. Thus full disturbabilty can be reached.

For the stochastic stabilization discussed in Section 5.5.3 full disturbability is an essential condition. Disturbability of position states is reached in the second measured epoch, velocity states starting from the third and acceleration states starting from the fourth measured epoch. State entries can only take part in the stochastic stabilization process which is essential for the adaptive filter once they are disturbable.

The delay or initialization lag implies that with only few epochs of observation data available the possibilities to correctly identify point velocities and especially point accelerations are limited.

5.5.5 Gain matrix K

The gain matrix $\mathbf{K}$ is an important term in the KF, as it defines the mean-square error minimum characteristic of the KF algorithm. It de-

termines the relationship between the innovations and the state vector estimate. In this section its structure is analyzed by means of the composing terms. As the detailed formulations are rather elaborate, they are given in Appendix B.

The gain matrix $\mathbf{K}_i$ is defined in (5.19).

$$\mathbf{K}_i = \mathbf{P}_{i|i-1}\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i)^{-1} = \mathbf{P}_{i|i-1}\mathbf{H}_i^T\mathbf{D}_i^{-1}. \quad (5.117)$$

Its structure is more closely analyzed by looking at the individual terms it is composed of, see Appendix B. Here the final result, which is based on the assumption $\mathbf{H}_{i\ vel} = \mathbf{H}_{i\ acc} = \mathbf{0}$, for all i is given:

$$\mathbf{K}_i = \begin{bmatrix} \mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T \\ \hline \mathbf{P}_{i|i-1\ vel|pos}\mathbf{H}_{i\ pos}^T \\ \hline \mathbf{P}_{i|i-1\ acc|pos}\mathbf{H}_{i\ pos}^T \end{bmatrix} \cdot (\mathbf{H}_{i\ pos}\mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T + \mathbf{R}_i)^{-1}. \quad (B.15)$$

The size of $\mathbf{K}_i$ is $[3u \times n]$, each element defines the relationship between one observation and one state vector entry. Only if state vector entries are $\neq 0$ this state gets update values $\neq 0$ in $\mathbf{K}_i\mathbf{d}_i$ and in $\hat{\mathbf{x}}_{i|i}$. In the first matrix product block in (B.15) it can be seen how entries in the gain matrix for states other than position are created: The off-diagonal blocks in the matrix $\mathbf{P}_{i|i-1}$, specifically $\mathbf{P}_{i|i-1\ vel|pos}$ create gain entries for the velocity states and $\mathbf{P}_{i|i-1\ acc|pos}$ for the acceleration states. Entries in these off-diagonal blocks are generated due to the structure of the system equation terms $\mathbf{\Phi}(i, i-1)$ and $\mathbf{\Gamma}(i-1)$. In the matrix term to be inverted in (B.15), (equivalent to $\mathbf{D}_i$ as defined in (5.22)), only position states are involved. Thus the error of the innovations $\mathbf{d}_i$ is exclusively influenced by the relationship between the position state components and the observations.

Overall, most of the various components resulting in the gain matrix $\mathbf{K}_i$ are predefined from the system description or the measurement equations. The only terms that can be actively influenced are the stochastic information terms: The initial state covariance matrix $\mathbf{P}_{0|0}$ is an approximation and is actively set at initialization stage and quickly looses its influence on the filter process. The two remaining terms are the

system noise covariance $\mathbf{Q}_{i-1}$ and the measurement covariance matrix $\mathbf{R}_i$. This property of the gain matrix structure is used in the stochastic stabilization procedure as described in Section 5.5.3. By acting on these terms the gain matrix can be changed in such a way that the KF adapts to misspecifications in the stochastic system noise and measurement models.

5.5.6 Selection of system model – Multiple-model filter

The KF system equations for the example of a geodetic deforming network are characterized by the degree of the development describing the equations of motion (i.e. if velocity and/or acceleration are included in the state). The system model has to be representative for the real situation, otherwise under- or over-parameterization may occur. In order to decide which model to choose, the KF itself can give the answer.

Multiple-model filters as discussed in e.g. [6] also provide adaptive filters: A discrete set for unknown parameters are introduced into a whole bank of KFs, working in parallel. Each filter gives a result and these results, weighted by the result's a posteriori probability, become combined into one estimate. The weights are determined from the innovations. A simplified and easily implemented method is encouraged in [41]: Instead of combining the various results of the whole filter bank, the one KF result corresponding to the smallest squared sum of innovations ($\sum \mathbf{d}_i^T \mathbf{D}_i \mathbf{d}_i$) is the final filter output. This is justified by the KF's property to minimize the expectation of the squared innovations, if the system model is correct (i.e. the optimal gain is found).

The approach presented in [41] is easily applied to the problem of finding the correct system model: Each point can in principle be described by three different models:

$$\mathbf{x}_j = \begin{bmatrix} \mathbf{x}_{j,pos} \\ \mathbf{x}_{j,vel} \\ \mathbf{x}_{j,acc} \end{bmatrix} = \begin{cases} \begin{bmatrix} \mathbf{x}_{j,pos} & \mathbf{0} & \mathbf{0} \end{bmatrix}^T & \text{or} \\ \begin{bmatrix} \mathbf{x}_{j,pos} & \mathbf{x}_{j,vel} & \mathbf{0} \end{bmatrix}^T & \text{or} \\ \begin{bmatrix} \mathbf{x}_{j,pos} & \mathbf{x}_{j,vel} & \mathbf{x}_{j,acc} \end{bmatrix}^T . \end{cases} \tag{5.118}$$

Even for small network sizes, all 3^u combinations (u being the number of network points) are not computable. If points which share some

kind of property and are expected to deform similarly can be summarized in groups, it is more effective to compute only the possible combinations for these groups. Groups can be built based on context information particular to the problem. In the case of a deforming cavern network, for example, all cavern floor points can be summarized to form one group.

Finding a suitable system model to match the real data is critical in each particular KF implementation. In some applications the necessary detail in the model description is suggested by the physical situation. In case of a geodetic cavern network, the point movements are considered to be very small and acceleration can generally be neglected.

5.6 Including additional information in KF

The result of a KF is a function of observation data. As much observation information as possible should be included. In many deformation applications some empirical knowledge about possible deformation behavior exists (e.g. points on a cavern floor are expected to move dominantly in vertical direction). Very often this information can not be quantified or modelled in a cause–response relation. Including this kind of additional information with a suitable stochastic model could support the KF estimation process, particularly in difficult observation configurations. An extension of the measurement model of the KF is presented in this section.

5.6.1 Formulating conditions as additional information

Including conditions or constraints on the estimation parameters in the parameter adjustment is a commonly used technique in geodetic applications, see for example [44, p. 177ff]. The conditions are formulated as 'virtual' observations. Linear observation equations are defined and added to the estimation process with an appropriate stochastic information.

A simple example of such a constraint is a *soft datum* definition, as already outlined in Chapter 4.1: As information serve the known coordinates of the datum points.

Applied to the KF algorithm the original measurement equation (5.7) is extended by the direct coordinate information (omitting the time index i):

$$\begin{bmatrix} \mathbf{z} \\ \mathbf{z}_d \end{bmatrix} = \begin{bmatrix} \mathbf{H}_b & \mathbf{H}_d \\ \mathbf{0} & \mathbf{I}_d \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_b \\ \mathbf{x}_d \end{bmatrix} + \begin{bmatrix} \mathbf{v} \\ \mathbf{v}_d \end{bmatrix} \tag{5.119}$$

where the top line represents the original measurement equation. The state vector is spilt up into a partition defining the datum $\mathbf{x}_d$ and a partition for the remaining states summarized in $\mathbf{x}_b$. The second line contains the direct coordinate information $\mathbf{z}_d$ and their direct relationship to the datum states. The stochastic model is extended to hold a covariance matrix for the direct coordinate information $\mathbf{R}_d$, no correlation is assumed with the measurement data:

$$\bar{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_d \end{bmatrix}. \tag{5.120}$$

Applied to the kinematic model implemented in the KF, the soft datum condition is imposed on the position states. Similar conditions can be formulated for the velocity and acceleration states, e.g. $\mathbf{x}_{vel} = \mathbf{x}_{acc} = \mathbf{0}$. Conditions on velocity and acceleration states change the structure of the measurement matrix $\mathbf{H}$, compare (5.67), as more matrix elements are assigned to values $\neq 0$:

$$\mathbf{H_i} = \begin{bmatrix} \mathbf{H}_{i,obs,pos} & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{i,add_obs,pos} & \mathbf{H}_{i,add_obs,vel} & \mathbf{H}_{i,add_obs,acc} \end{bmatrix}. \tag{5.121}$$

In deformation analysis applications additional information about the expected point movements can be introduced. The different types of information depend on the example at hand, here a few possibilities are listed:

- Direction of movement imposed on velocity components: A point is moving in a certain direction. Only the direction is constrained but not the amount. Thus two additional information equations are defined, for the local azimuth θ and zenithal angle ϕ. Both expressions are non-linear:

$$x_{vel,j} - \cos\theta \sqrt{x_{vel,j}^2 + y_{vel,j}^2} = 0 \tag{5.122}$$

$$z_{vel,j} - \cos\phi\sqrt{x_{vel,j}^2 + y_{vel,j}^2 + z_{vel,j}^2} = 0 \tag{5.123}$$

- Movement along single coordinate direction imposed on velocity components: If point movements are assumed to be only along one coordinate direction, (5.122) and (5.123) can be replaced by constraining the other coordinate directions to 0 (e.g. movement expected in $\pm z$, $\mathbf{x}_{vel/acc} = \mathbf{y}_{vel/acc} = \mathbf{0}$ are constrained), resulting in linear expressions.

- Relative constraints can model relationships between points, e.g. same point movements or movement in the same direction. This can be extended to model more complex relationships.

In general the additional information equations are not linear and have to be linearized. Good approximation values are needed for the linearization. If this is not the case, the possibilities of including additional information are very limited.

Including additional information as outlined above in an adaptive KF for a geodetic network in kinematic setup can help to support the estimation solution, especially if the measurement configurations are incomplete. It can be of benefit to stabilize the estimation solutions by constraining the states for points missing in one or several epochs.

Results for the application to a practical example are presented in Chapter 6.2.

5.7 Conclusion

In this chapter an adaptive KF in a kinematic setup to model and adjust a 3D geodetic network has been presented. The application of this approach has been motivated for the problem of handling a network of geodetic measurements in changing configuration that is subject to deformation. The KF algorithm terms in discrete formulation as commonly used in literature have been derived for this case.

The adaptive KF has been developed on the basis of stochastic stabilization. The stochastic models of the system noise and measurement data are modified iteratively in order to react to unexpected point movements and gross errors in the observation data. The structure of the

resulting algorithm and related variables are analyzed in more detail and some suggestions given to further enhance the algorithmic performance in the application to practical examples.

6

Application and results for an implementation of the adaptive Kalman Filter

The adaptive KF algorithm presented in Chapter 5 is applied to data to show its capabilities, performance and possible shortcomings. Both simulated and real survey data are processed and the results are presented in this chapter, which is divided in four sections: In Section 6.1 the implementation of the algorithm is shortly described. In Section 6.2 results for the application to simulated data for a small network example are given to show the potential performance of the algorithm and the influence of stochastic stabilization. The benefit of including additional information about expected point movements is demonstrated. In Section 6.3 data for the theoretical ATLAS cavern network configuration is simulated and results of the KF algorithm's application presented. Conclusions about the major difficulties of the network layout are drawn and how the KF approach helps to overcome these problems. Finally in Sec-

tion 6.4 results are given for the algorithm's application to real ATLAS cavern network data. Data gathered over a period of three years are analyzed.

The results are presented by plots and tables of estimation results, point movements (deformation) and corresponding error estimates. Comparisons to conventional least squares solutions are used to discuss the superiority but also limitations of the adaptive KF algorithm.

6.1 Implementation of the adaptive KF algorithm

The adaptive KF algorithm that models a geodetic network in a kinematic system as described in Chapter 5.5 is implemented in MATLAB©6.5. The resulting application reads survey data in the standard input format (*.inp) of CERN's least squares adjustment software package LGC, [31]. The implementation includes all necessary least squares setup processes. Due to the complex program structure of LGC it was not possible nor encouraged to develop and implement the adaptive KF filter algorithm interacting directly with the corresponding modules in LGC.

Point and measurement declarations that are implemented in the MATLAB application include the following (abbreviations used refer to LGC nomenclature):

Point declarations:

CALA: Fixed datum points
POIN: Unknown network points

In the KF implementation **CALA** points are set up as static, fixed 3D points. Point velocity and acceleration are set to 0 and point position, velocity and acceleration coordinate components are not allowed to change. All other points are summarized in **POIN** and are set up as full 3D kinematic points. Further LGC point declaration options (e.g. **VZ** for points variable only in z and **VXY**, **VXZ** or **VYZ** accordingly) are interpreted as **POIN**.

Observation declarations:

DMES:	3D distance
DHOR:	2D horizontal distance
DTHE:	3D distance between theodolite and target (instrument height unknown)
DVER:	Vertical distance or height difference (geometrical levelling)
ANGL:	Horizontal angle
ZENI:	Vertical angle (instrument height unknown)
ZENH:	Vertical angle (instrument height known)
ECHO:	Horizontal offset measurements

The implemented processing scheme follows the scheme derived for the adaptive KF algorithm in Chapter 5 which is summarized in Figure 5.3. Some more practical details are listed in the following:

- *Model definition:* The 'degree' of polynomial for the kinematic model is set a priori, i.e. include point velocities and accelerations or only velocities in the state vector. This definition can be made for every single point. The decision for the right model can be made applying the method of multiple-model filters that has been shortly outlined in Chapter 5.5.6.
- *Initialization:* The state vector and its variance are initialized according to (5.62). Approximate values for the point coordinates are taken from the points' first occurrence in a data input file. The point error information is chosen a priori, e.g. 5 mm for an individual point coordinate. In [6] it is suggested to use not too large values, as this can cause numerical problems resulting in filter divergence.
  The initial system noise errors are chosen relatively small (e.g. $0.005\,\mathrm{mm}/dt^3$), as these values are iteratively adapted if necessary. A small initial system noise error keeps the uncertainty about the system description to a minimum. If the initial value would be chosen too large, the system description would be unnecessarily degraded each time it is predicted to the next measurement epoch. The optimum value for a certain data set (defining optimal by e.g. minimizing the squared sum of innovations $\mathbf{d}^T\mathbf{D}^{-1}\mathbf{d}$) can be found by using the method of multiple-model filters for different system noise variance values.

- *Prediction:* The state estimate and its variance are projected to the next epoch of available measurements using the transition matrix as defined in Chapter 5.5.1. If it is the first epoch after the initialization step the transition is applied with a time difference $\Delta t = 0$.

- The *measurement matrix* $\mathbf{H}_i$ is set up according to (5.67) using the linearized measurement equations. The implemented types of measurements are limited to the group listed above.

- The *gain matrix* is determined as in (5.19) including the numerically demanding inversion of the innovation variance matrix $\mathbf{D}_i$.

- A *preliminary update* for the state vector and its variance matrix is derived, according to (5.66) and (5.11).

- The *global model test* as defined in (5.76) decides whether a significant misspecification is indicated by the innovations.

- If the global model test indicates an outlier either in the system description or in the observation data, it is localized and the corresponding stochastic information changed according to the iterative procedure of *stochastic stabilization* described in Chapter 5.5.3. An outlier in the system description is an unexpected change in the system model (e.g. unexpected point deformation). By *increasing the system noise variance* the uncertainty in the system description is increased. This allows for more freedom in the system model estimation and more weight is given to the measurement data to derive coordinate estimates. In the case of an outlying observation the corresponding observation error is increased in order to reduce this observation's influence on the estimation result. The process of stochastic stabilization is carried out iteratively in each epoch of observation data which enables a fast and efficient reaction of the KF to new information.

- For the global model test and the derivation of the test statistic for the stochastic stabilization the chosen level of significance is 95%. The disturbability of the state vector entries (which indicate misspecifications in the system description) needs to be established beforehand, as only disturbable states are to be considered in the derivation of the test statistic for stochastic stabilization. Thus

Table 6.2: *Small simulated network – Statistics of the individual data sets: Epoch number, time, number of fixed datum points (CALA), number of variable network points (POIN), number of measured spatial distances (DMES), number of measured vertical distances (DVER), number of measured horizontal angles (ANGL), number of vertical angles (ZENI), their respective mean redundancy numbers ($\bar{r}_i$) and the overall mean redundancy number.*

Epoch	Time	CALA	POIN	DMES		DVER		ANGL		ZENI		Mean
	[]	N°	N°	N°	$\bar{r}_i$	N°	$\bar{r}_i$	N°	$\bar{r}_i$	N°	$\bar{r}_i$	redund.
0	0	2	10	84	0.80	19	0.54	84	0.96	84	0.89	0.89
1	1	2	10	28	0.44	19	0.49	28	0.85	28	0.71	0.71
2	2	2	10	24	0.33	19	0.49	24	0.85	24	0.67	0.67
3	3	2	10	28	0.53	19	0.49	28	0.91	28	0.75	0.75
4	4	2	8	28	0.49	0	—	28	0.69	28	0.71	0.71
5	5	2	10	27	0.52	19	0.49	27	0.89	27	0.74	0.74
6	6	2	14	55	0.62	23	0.55	55	0.87	55	0.78	0.78
7	7	2	18	43	0.54	29	0.50	43	0.81	43	0.71	0.71
8	8	2	16	58	0.56	0	—	58	0.70	58	0.72	0.72
9	9	2	18	118	0.76	29	0.65	118	0.93	118	0.86	0.86

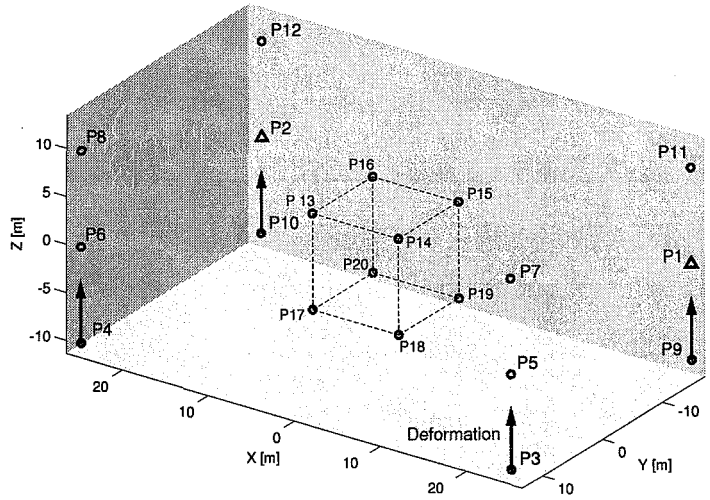


Figure 6.1: *Small simulated network – Layout: Datum points P1 and P2, network points P3–P12 and object points P13–20. Deformation acting on four 'cavern floor' points P3–P4 and P9–P10.*

Table 6.1: *Small simulated network: Key data set parameters.*

Data source	Simulation
N° of points	20
datum	2
N° of moving points	
assumed	18
true	4
N° of measurement epochs	10
3D	all
Data sampling	1

The network is assumed to be deforming in the four 'floor' points: A rise of all four points with a velocity of 0.1 mm/dt and an acceleration of $0.02\,\mathrm{mm/dt^2}$ has been applied.

See Table 6.1 and Table 6.2 for a summary of some key parameters of the data set. The redundancy numbers in Table 6.2 refer to the individual (SE) LSQ solutions and are not the KF redundancy numbers.

The adaptive KF implementation is applied to this data set. All points except the datum points are set up as kinematic points.

Some additional options are available in the implemented application:

- Possibility to include additional information about e.g. expected point movements as described in Chapter 5.6.
- Declaration of temporary or discontinued points by adding '#' at the end of the point declaration in the input data file.
- Including deterministic information about point position changes in additional data input files (*.detinp). This can be useful if, for example, a point monument is changed, e.g. socket without reference sphere is changed to standard reference socket with sphere (for illustration see Figure 3.4, where the vertical offset is 44.45 mm.
- Declaration of a soft datum as described in Chapter 4.1.1. Points defined as **CALA** in the input data file are interpreted as soft datum definition.
- Several plotting options are available.

6.2 Small simulated network data

A small network in the approximate dimensions and simplified design of the ATLAS cavern network is simulated, represented by twelve network points (P1–P12) and eight object points (P13–P20), corresponding to possible detector object points. The datum definition is realized by two fixed network points (P1 and P2) inside the cavern, see Figure 6.1.

Measurement data in ten epochs ($\Delta t = 1$) is simulated in different configurations, including horizontal and vertical angle measurements ($\sigma = 7^{cc}/10^{cc}$), distance measurements ($\sigma = 0.2$ mm/0.3 mm) and some vertical distance measurements ($\sigma = 0.1$ mm/0.3 mm). Different assumed measurement precisions represent different measurement conditions e.g. one– or two–face measurements or different target options. The data has been simulated in a small program implemented in MATLAB©6.5.

stochastic stabilization for the system description can only start in the second epoch of available data.

- State vector entries referring to one coordinate component of one point are grouped for the localization step (i.e. x-position, x-velocity and – if considered – x-acceleration state entries for point j form one group). No grouping of horizontal coordinate components is applied, as the point deformations are expected to be dominantly along the individual coordinate directions, i.e. radial and vertical.

- The *detection of outliers* in the observation data starts in the very first epoch. The process from the prediction step onwards is iterated with the modified stochastic information for the system noise and for the measurements, until the global model test passes.

- Statistical test terms are derived based on the adapted level of significance, as suggested in (5.108). This enables to make the test more sensitive in the case of poorly controlled observations.

- If the reliability number for an observation is zero it is excluded from the gross error detection. Such observations (e.g. HLS measurements with no direct link to the full network) containing information new to the system description would be completely downweighted in the estimation process instead of adapting the system description.

- The update of the state vector and its variance is *valid* if the global model test does not find any significant innovation. The corresponding variables are *stored* in a flexible *data structure.*

- For the next epoch of measurement data the process starts again at the prediction step. The stochastic information for the system noise is reset to its initial values. Points that are already known to the system are projected by applying the transition matrix. Points that are new to the system are initialized with approximate coordinate values and a chosen a priori variance information as in the initialization step at the very beginning of the filter algorithm.

This small and clear example shows the potential performance of the adaptive KF algorithm in kinematic setup. Results for two points — one moving network point and one stationary object point — are presented by means of plots and tables summarizing point position estimates and corresponding error information. The KF results are compared to true positions and the results of individual, independent least-squares adjustments, from now on referred to as *Single Epoch* (SE) Least-SQuares (LSQ) results. The SE LSQ algorithm is also implemented in MATLAB©6.5 and refers to the same basic adjustment functions as the KF implementation (i.e. setup of linearized observation equations etc.). The approximate point coordinates used to derive the SE LSQ solution are the true simulated point positions which are not available to the KF after the initial epoch. This gives an advantage to the SE LSQ results over the KF results.

Two system models are set up for this example: The first system model (SMI) includes point position and velocity components and in the second (SMII) also point accelerations are estimated. Results are presented for both models and a multiple–model filter as outlined in Chapter 5.5.6 is applied to decide on the better suited system model.

Three examples of active stochastic stabilization as presented in Chapter 5.5.2 show the influence of the adaptive process on the Kalman gain matrix. This matrix determines the influence of observations on the system state vector. Changes in the stochastic information for the measurement data and the system noise induce changes in the gain matrix. The analysis of such changes give valuable insight into the adaptive KF process.

Additional information on expected point movements is included as formulated in Chapter 5.6 and the results are presented.

6.2.1 Network configuration and observations

Observations for ten epochs are simulated. The configuration is different in all epochs. In epochs 1–6 only network points are present, from epoch 7 onwards also object points are measured. Object points are target points and never instrument stations, consequently no reciprocal measurements are available for this group of points.

An example of one epoch of simulated data can be seen in Figure 6.2. In this epoch 9 most of the network points and all of the object points

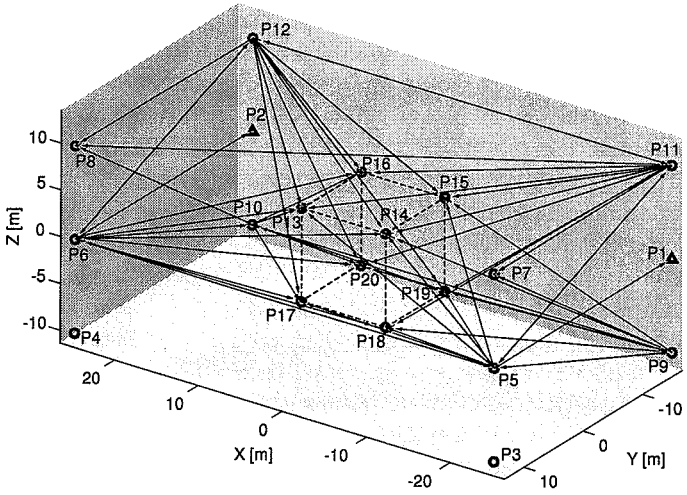


Figure 6.2: *Small simulated network: Simulated observations in epoch 9. Tachymetric measurements from network points P5, P6 and P9–P12.*

are observed, with theodolite stations on network points P5, P6 and P9–P12. Measurements include horizontal and vertical angles as well as spatial distances. The datum points P1 and P2 are only sighted from one station each, thus the reference to the datum is weak, as these observations are poorly controlled.

Network configurations in other epochs are similar but all network points are measured at least in nine epochs, object points are measured in three or four epochs. The datum is always included in the measurements.

6.2.2 Analysis and comparison of adaptive KF results with single epoch LSQ results and true values

The results for the adaptive KF in kinematic setup are compared to the SE LSQ results. The subsequent plots show the results in xy- and yz-plane. An estimated position is indicated by '+' for the SE LSQ result and '×' for the KF result. The corresponding error information is represented by the *Helmert error ellipse* (1σ) describing an area of confidence

of approximately 39%, [44, p. 259]. Confidence ellipses representing a significance level of 95% are not plotted because these objects would be too large for illustration of the position results. As this is a simulated example, also the true positions are known and are plotted as '□'. In overview plots showing the SE LSQ results separately from the KF results for all epochs, e.g. Figure 6.3, a color scheme is used to show the different epochs in time: Earlier epochs are plotted in lighter shades of grey, getting darker until black for the most recent epoch. Deformation vectors (i.e. velocity state estimates) are plotted for the KF results. If no observation is available and the KF estimate is the predicted value only, the corresponding error ellipse is plotted in a dashed line. More detailed plots give the results for individual epochs separately, giving a better overview of the results with respect to time. Numeric results are summarized in tables.

The results for two points are presented: One moving network point (P10) and one stationary object point (P18). Results for the System Model I (SMI) including only point velocities and for the System Model II (SMII) including point velocities and accelerations are shown.

Moving network point P10	
Measured in	9 epochs
3D	9 epochs (time 0, 1, 2, 3, 5, 6, 7, 8, 9)
levelling only	none

The moving network point P10 has been observed in nine out of ten epochs. It is only moving in the z-direction.The SE LSQ results in the xy-plane in Figure 6.3 show a random distribution around the true position, the offsets from the true value are generally not significant with respect to the 1σ- error ellipses corresponding to a level of significance of 39%, except in one case.

The results for the SMI in the xy-plane in the middle of Figure 6.3 show that overall no significant point movement in x- or y-direction can be identified, which is correct. The movement is illustrated by the estimated deformation or velocity vector. To indicate a significant movement this vector would have to extend clearly beyond the 1σ-error ellipse corresponding to the 39% level of significance.

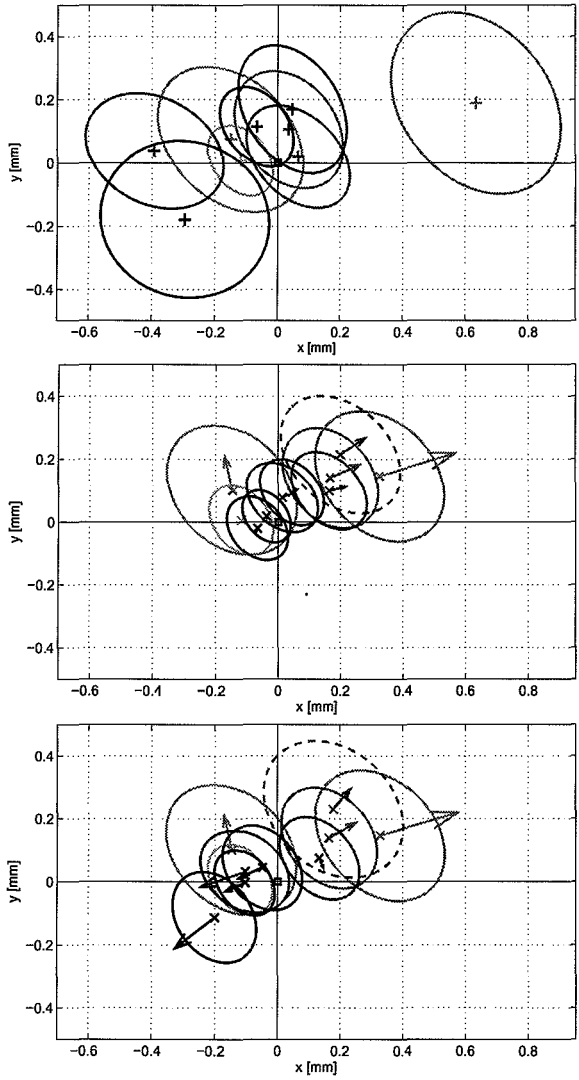


Figure 6.3: *Small simulated network – Moving network point P10 (movement in z-direction only): Comparison of SE LSQ results (top), KF results for system model including point velocities (SMI) (middle) and KF results for system model including point velocities and accelerations (SMII) (bottom) in the xy-plane.*

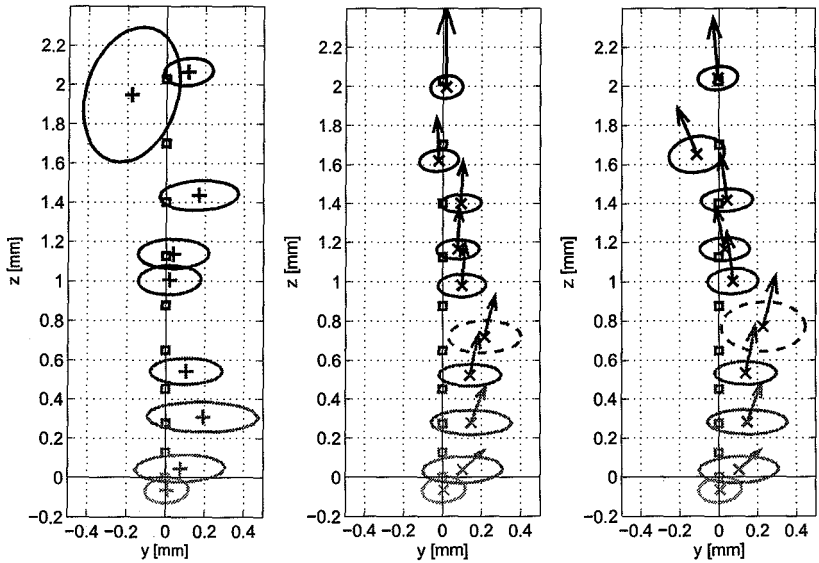


Figure 6.4: *Small simulated network – Moving network point P10: Comparison SE LSQ results (left), KF results for system model including point velocities (SMI) (middle) and KF results for system model including point velocities and accelerations (SMII)(right) in the yz-plane.*

In case of the KF, generally the error estimations decrease in later epochs as more and more data is available to derive the point and error estimates. In epoch 2 the point is not so well determined in x-direction as it is only observed in 3D from station P3 and P4, which are also moving network points. Its vertical position is determined by geometrical levelling. The SE LSQ result is more affected by this bad configuration than the KF result as in this case only data from this very measurement epoch can be used, the estimate is approximately 0.63 mm off in x-direction and 0.19 mm in y-direction, see also Table 6.3.

Comparing the results for the SMI including only point positions and velocities to the results for the SMII including additionally point accelerations it is seen that in both cases no significant horizontal point movement can be identified. The KF with the SMI estimates zero velocity values after some epochs but the KF with the SMII derives deformation vectors in changing directions also in later epochs. It is more difficult for the higher parameter model to identify zero movement.

Table 6.3: *Small simulated network – Moving network point P10: True simulated coordinates and results for SE LSQ (top) and KF (SMI) (bottom): Differences to true starting position in epoch 0.*

Epoch	True positions [mm]			SE LSQ results [mm]					
	x	y	z	x ± rms		y ± rms		z ± rms	
0	0.00	0.00	0.00	-0.12	0.10	0.01	0.11	-0.07	0.06
1	0.00	0.00	0.13	-0.15	0.23	0.07	0.23	0.05	0.07
2	0.00	0.00	0.28	0.63	0.27	0.19	0.29	0.31	0.08
3	0.00	0.00	0.45	0.04	0.17	0.11	0.19	0.54	0.07
4	0.00	0.00	0.65	—	—	—	—	—	—
5	0.00	0.00	0.88	0.07	0.17	0.02	0.16	1.00	0.07
6	0.00	0.00	1.13	-0.39	0.22	0.04	0.18	1.14	0.07
7	0.00	0.00	1.40	0.05	0.17	0.17	0.20	1.44	0.08
8	0.00	0.00	1.70	-0.29	0.27	-0.18	0.25	1.95	0.34
9	0.00	0.00	2.03	-0.07	0.12	0.11	0.13	2.06	0.07

Epoch	KF positions [mm]						KF velocities [mm/dt]					
	x ± rms		y ± rms		z ± rms		vx ± rms		vy ± rms		vz ± rms	
0	-0.12	0.10	0.01	0.11	-0.07	0.06	0.00	0.50	0.00	0.50	0.00	0.50
1	-0.15	0.21	0.10	0.21	0.04	0.07	-0.02	0.22	0.11	0.22	0.10	0.09
2	0.33	0.21	0.15	0.21	0.28	0.06	0.24	0.12	0.07	0.12	0.18	0.05
3	0.17	0.15	0.14	0.16	0.52	0.06	0.10	0.06	0.04	0.07	0.21	0.03
4	0.20	0.19	0.22	0.19	0.72	0.08	0.08	0.06	0.05	0.06	0.20	0.03
5	0.16	0.12	0.10	0.12	0.98	0.06	0.05	0.03	0.01	0.03	0.22	0.02
6	0.01	0.12	0.08	0.11	1.16	0.05	0.01	0.03	0.01	0.03	0.21	0.01
7	0.04	0.10	0.09	0.10	1.40	0.05	0.01	0.02	0.01	0.02	0.21	0.01
8	-0.07	0.10	-0.02	0.10	1.62	0.05	0.00	0.02	-0.01	0.02	0.22	0.01
9	-0.04	0.08	0.02	0.08	2.00	0.06	0.00	0.02	0.00	0.02	0.41	0.13

The results for the yz-plane in Figure 6.4 show that the vertical point movement is easily identified by both the SE LSQ and the KF solutions. In case of the KF position estimates are obtained even though no observations are available (as in epoch 4). The corresponding error ellipses are slightly inflated which is reasonable, as the corresponding position estimates are only predictions and the system descriptions are subject to uncertainty.

In epoch 8, the network configuration is degraded, as the connection to the datum points is only very weak (see Figure 6.2). The redundancy numbers for the corresponding observations are very small (e.g. $r_i < 0.02$ for the spatial distance measurements). The SE LSQ result differs notably from the true position while the KF is generally only slightly affected. The KF at this stage can rely on a good system description to compensate for poorly controlled observations while the SE LSQ estimate is affected by any errors in these uncontrolled observations decreasing the overall network reliability. The clear acceleration in the movement should be better followed with the SMII model. The improvement is in fact small as can be seen in the plot, however the error estimation for the position is larger due to the increased number of unknowns. Also the estimation is more affected by missing observations or bad measurement configuration.

Using the indicator min $(\sum \mathbf{d}_i^T \mathbf{D}_i \mathbf{d}_i)$ as used in the multiple–model filter outlined in Chapter 5.5.6 to decide on the better suited system model, the decision is by a marginal difference in favor for the SMII (with a value for $\sum \mathbf{d}_i^T \mathbf{D}_i \mathbf{d}_i = 1597.6$ against the SMI with $\sum \mathbf{d}_i^T \mathbf{D}_i \mathbf{d}_i = 1673.4$.

Overall the improvement due to the higher–parameterized system model including point accelerations is small although the acceleration in this example is clearly identifiable in the data. For this reason the results in the following are only shown for the application of the SMI. Numerical details for the SE LSQ results and the KF (SMI) results can be found in Table 6.3. The *root mean square* (rms) errors of each estimated value are given.

In a more sophisticated solution the multiple–model filter would be set up for all combinations of system models considering point and coordinate components individually. This demands for considerable computing power even for a small example and becomes impracticable for larger networks.

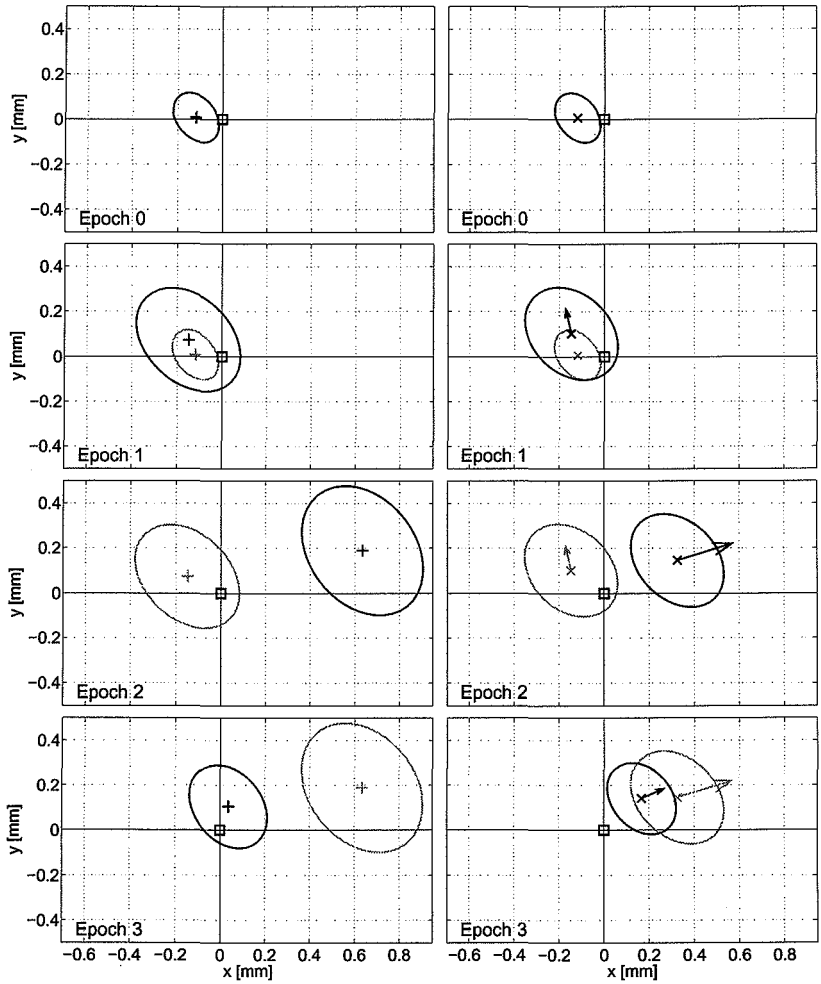


Figure 6.5: *Small simulated network – Moving network point P10 (movement in z direction only), Epoch 0 – Epoch 3: SE LSQ results (left column) and KF (SMI) results (right column) in the xy-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

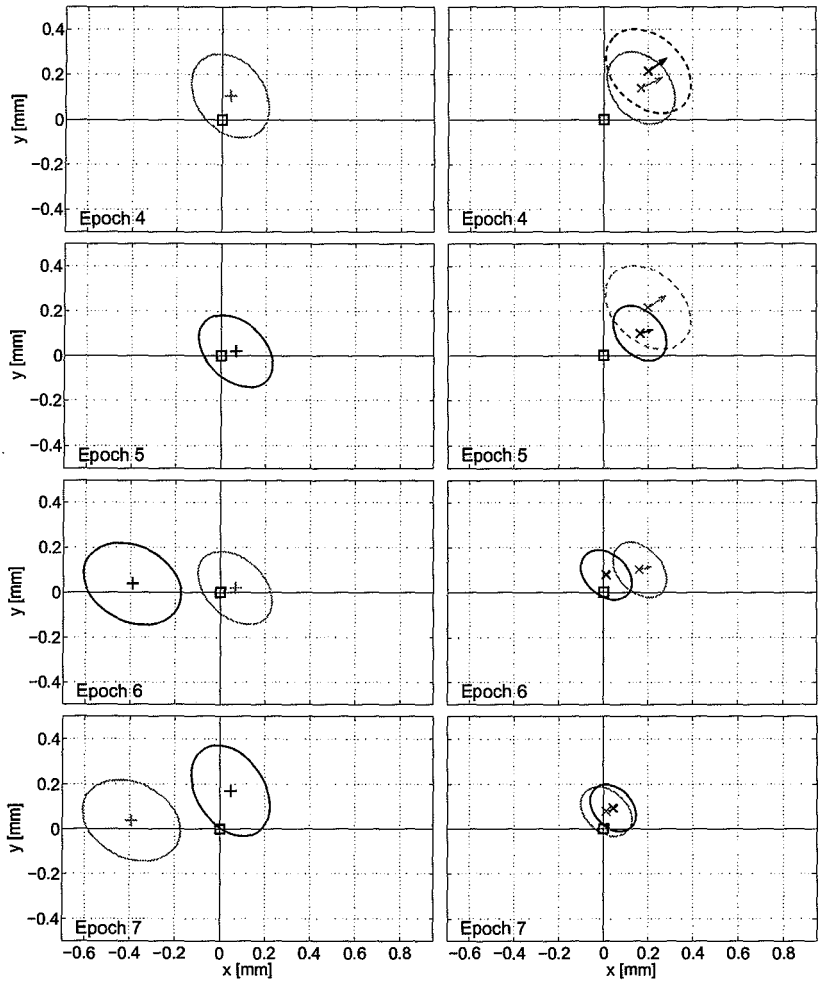


Figure 6.6: *Small simulated network – Moving network point P10 (movement in z direction only), Epoch 4 – Epoch 7: SE LSQ results (left column) and KF (SMI) results (right column) in the xy-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

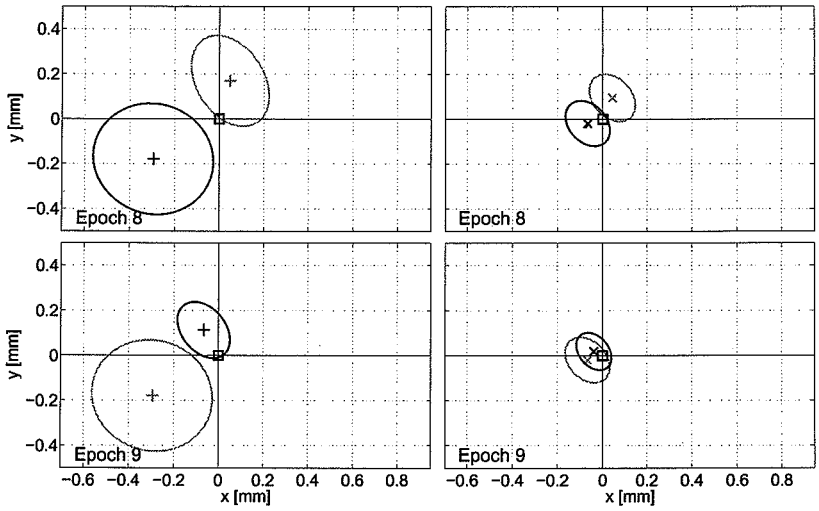


Figure 6.7: *Small simulated network – Moving network point P10 (movement in z direction only), Epoch 8 – Epoch 9: SE LSQ results (left column) and KF (SMI) results (right column) in the xy-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

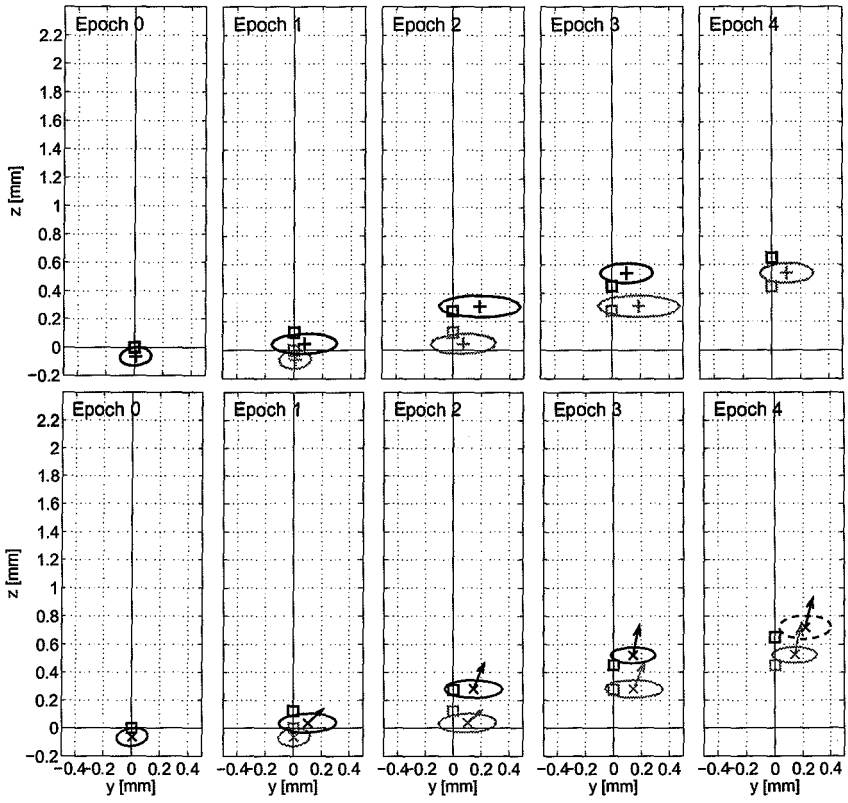


Figure 6.8: *Small simulated network – Moving network point P10 (movement in z direction only), Epoch 0 – Epoch 4: SE LSQ results (top row) and KF (SMI) results (bottom row) in the yz-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

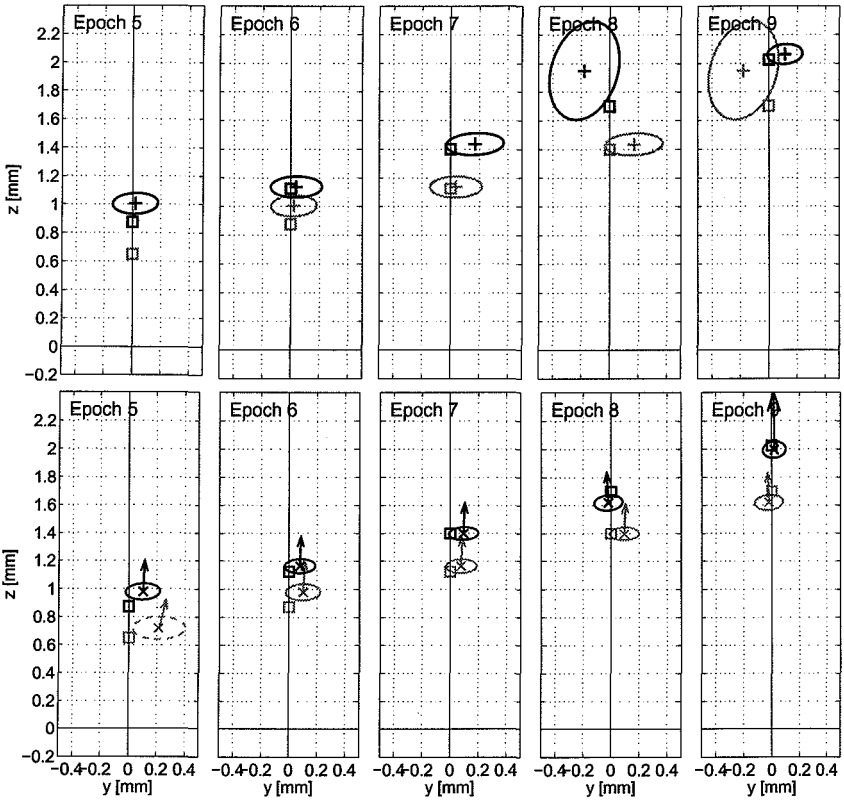


Figure 6.9: *Small simulated network – Moving network point P10 (movement in z direction only), Epoch 5 – Epoch 9: SE LSQ results (top row) and KF (SMI) results (bottom row) in the yz-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

For a better visualization of the position results with respect to time may serve Figure 6.5 – Figure 6.9 illustrating the SE LSQ and KF SMI results as in Figure 6.3 and Figure 6.4 but epoch-by-epoch: The SE LSQ ('+') and KF ('×') results for each epoch are plotted separately and arranged next to each other for better comparison. The estimates from the previous epoch are plotted in grey lines.

Stationary object point P18	
Measured in	4 epochs
3D	3 epochs (time 6, 8, 9)
levelling only	1 epoch (time 7)

The object point P18 has been measured in the last four epochs, in epoch 7 only by geometrical levelling. The illustrations of the results in the xy-plane in Figure 6.10 and for the yz-plane in Figure 6.11 show that the KF correctly identifies this point as being stationary.

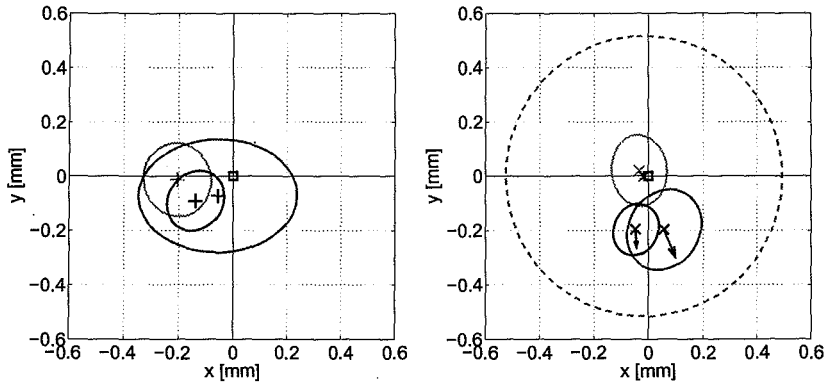


Figure 6.10: *Small simulated network – Stationary object point P18: Comparison of SE LSQ results (left) and KF (SMI) results (right) in the xy-plane.*

The KF results are for the simpler system model SMI. Detected point movements are not significant for a level higher than the 39% (1σ) confidence ellipse.

The results in epoch 7 in Figure 6.10 show a largely inflated error ellipse. This is understandable as in this epoch no horizontal measurement information is available and the estimate is derived from only one

Table 6.4: *Small simulated network – Stationary object point P18: True simulated coordinates and results for SE LSQ (top) and KF (SMI)(bottom): Differences to true starting position.*

Epoch	True positions [mm]			SE LSQ results [mm]					
	x	y	z	x ± rms		y ± rms		z ± rms	
0	0.00	0.00	0.00	—	—	—	—	—	—
1	0.00	0.00	0.00	—	—	—	—	—	—
2	0.00	0.00	0.00	—	—	—	—	—	—
3	0.00	0.00	0.00	—	—	—	—	—	—
4	0.00	0.00	0.00	—	—	—	—	—	—
5	0.00	0.00	0.00	—	—	—	—	—	—
6	0.00	0.00	0.00	-0.20	0.12	-0.01	0.14	-0.20	0.14
7	0.00	0.00	0.00	—	—	—	—	-0.18	0.08
8	0.00	0.00	0.00	-0.05	0.29	-0.07	0.21	0.23	0.28
9	0.00	0.00	0.00	-0.14	0.11	-0.09	0.11	0.08	0.11

Epoch	KF positions [mm]						KF velocities [mm/month]					
	x ± rms		y ± rms		z ± rms		vx ± rms		vy ± rms		vz ± rms	
0	—	—	—	—	—	—	—	—	—	—	—	—
1	—	—	—	—	—	—	—	—	—	—	—	—
2	—	—	—	—	—	—	—	—	—	—	—	—
3	—	—	—	—	—	—	—	—	—	—	—	—
4	—	—	—	—	—	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—	—	—	—	—
6	-0.03	0.10	0.02	0.13	-0.20	0.13	0.00	0.50	0.00	0.50	0.00	0.50
7	-0.02	0.51	0.00	0.52	-0.16	0.22	0.00	0.50	0.00	0.50	0.02	0.24
8	0.06	0.14	-0.20	0.15	0.08	0.15	0.04	0.08	-0.10	0.09	0.14	0.10
9	-0.05	0.08	-0.20	0.09	0.08	0.09	0.00	0.04	-0.06	0.05	0.07	0.05

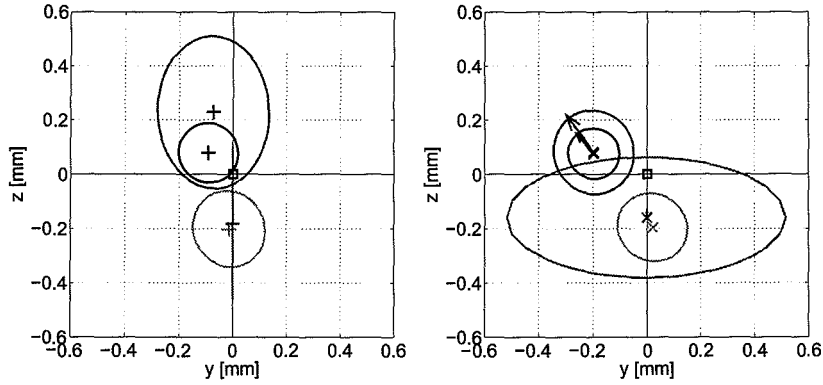


Figure 6.11: *Small simulated network – Stationary object point P18: Comparison of SE LSQ results (left) and KF (SMI) results (right) in the yz-plane.*

previously measured epoch. In Figure 6.11 the SE LSQ estimate for this epoch 7 is the levelling result only, the error information is represented by an error bar. The corresponding error figure for the KF result is an ellipse plotted in a full line, as the z-position is derived as a state update using observation but the y-component of the position is only a prediction. See Figure 6.12 and Figure 6.13 for the visualization of the results with respect to individual epochs.

Table 6.4 lists the numerical results for this point.

In Table 6.2 the overall mean redundancy numbers have been given for all individual measurement epochs. For the same data also mean values of the KF observation redundancy numbers as defined in (5.107) can be derived.

The comparison of these numbers gives a measure

Table 6.5: *Small simulated network – Overall mean redundancy numbers for SE LSQ and KF (SMI) solutions.*

Epoch	Mean redundancy		Reliability
	SE LSQ	KF	gain in %
0	0.89	0.89	0
1	0.71	0.74	4
2	0.67	0.75	11
3	0.75	0.82	9
4	0.71	0.87	18
5	0.74	0.84	12
6	0.78	0.85	9
7	0.71	0.83	14
8	0.72	0.83	13
9	0.86	0.90	5

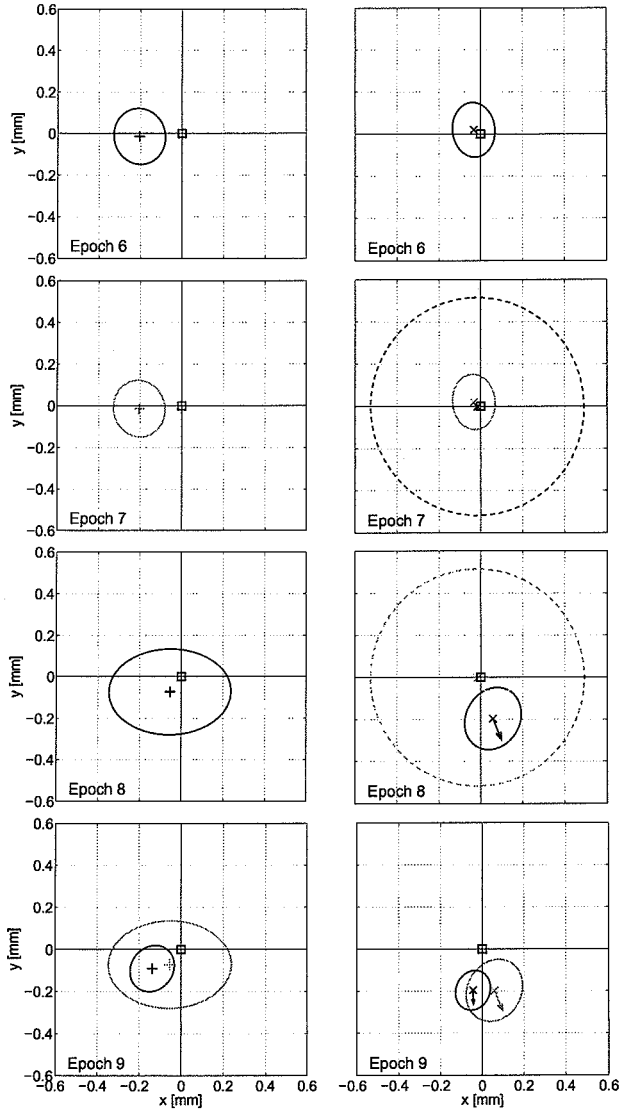


Figure 6.12: *Small simulated network – Stationary object point P18: Comparison of SE LSQ results (left column) and KF (SMI) results (right column) in the xy-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

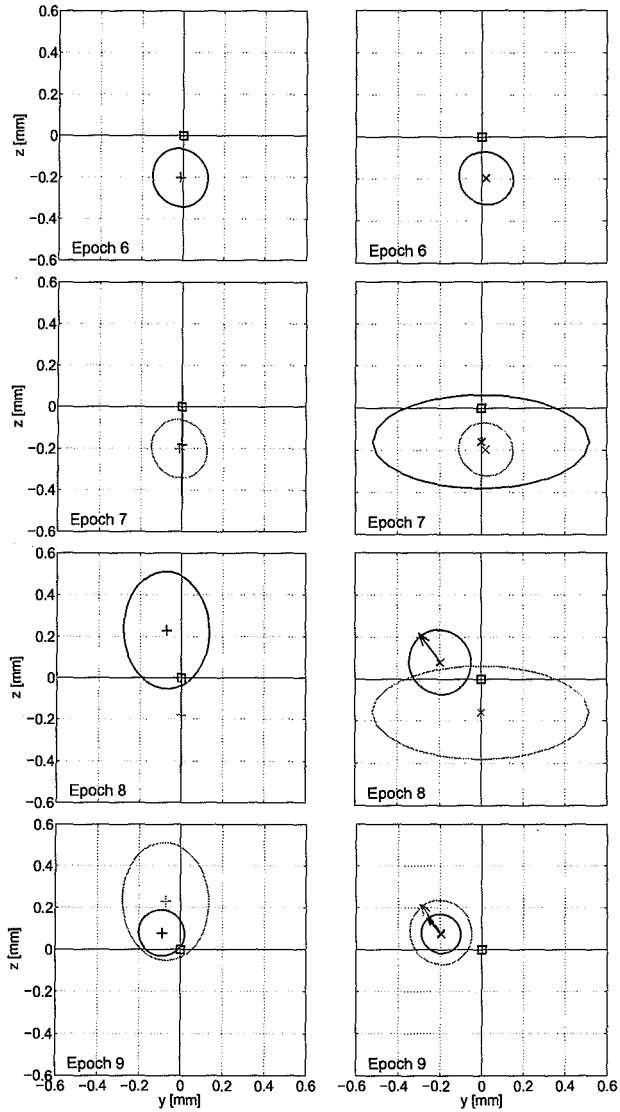


Figure 6.13: *Small simulated network – Stationary object point P18: Comparison of SE LSQ results (left column) and KF (SMI) results (right column) in the yz-plane for each epoch. Estimates from previous epoch plotted in grey lines.*

for the increase in network reliability if the KF is applied, see Table 6.5. The gain compared to the SE LSQ results is larger (between 13 – 18%) if the individual epoch measurement configurations are bad, see for example epoch 4, 7 or 8. If the individual measurements are well controlled (as in epoch 9) the gain is smaller (5%).

6.2.3 Analysis of the influence of stochastic stabilization on the KF gain matrix

In this section the effect of the stochastic stabilization (see Chapter 5.5.3) on the gain matrix **K** is discussed. In Appendix B the structure of the gain matrix is discussed in analytical detail. To better understand how it is influenced in the adaptive filter due to the stochastic stabilization process, some details in the context of this small application example are illustrated.

The adaptive KF identifies discrepancies between the system model and the current measurements by means of innovation analysis. Discrepancies can be caused by misspecifications in the system model or gross errors in the observation data. In the example of a kinematic network, misspecifications in the system model can be caused by unexpectedly large deformations. The system model is adapted by iteratively increasing the system noise error until the system description and the measurement data agree. Gross errors in the observation data also cause discrepancies between the system model and the measurement data. It has been shown in Chapter 5.5.3 that the system noise stochastic model can be adapted in parallel to the stochastic model for the observations in order to decrease the influence of outlying observations.

Changes in the stochastic models of the system noise and the measurement data influence the gain matrix. Using the example of the simple data set discussed in this section and selecting three particular cases of stochastic stabilization, the effects on the gain matrix are illustrated. The three cases include two deliberately added gross errors in one poorly and one well controlled observation. The effect of adapting the system description by increasing the system noise error is demonstrated for one state parameter. This example uses the original data set analyzed in this section.

The gain matrix $\mathbf{K}$ describes the influence of individual observations on individual state vector entries. It has a dimension of $[u \times n]$ where u describes the number of unknowns (in the case of the SMI u is $6 \times$ the number of points) and n describes the number of observations.

Gross error detection

In each of the following examples one single gross error is added deliberately to the original data set. The two examples are not connected, i.e. only one gross error is present in the data set each time.

The gain matrix is initially not affected by the deliberate introduction of a gross error, as it is by definition not dependent on the observation data. The immediate effect of a gross error is on the innovation $\mathbf{d}$ and thus the predicted state vector residuals $\mathbf{v}_{\mathbf{x},i}$ and the measurement residuals $\mathbf{v}_{\mathbf{z},i}$, as defined in (5.70).

The gain matrix is affected by changes to the stochastic models, as discussed in Chapter 5.5.5. In the case of a gross error the measurement error model is adapted by the algorithm in order to reduce the effect of the outlying observation.

Misspecifications indicated by the global model test are localized and the corresponding error information, i.e. measurement errors, increased. Changes to the stochastic model affect the gain matrix entries in the following iteration adjustment step. The iteration is continued until the global model test passes. Only then can the data processing go ahead to include the next epoch of measurement data.

Figure 6.14 illustrates how the effect of the change in the measurement error on the gain matrix can be visualized. The entries in the gain matrix $\mathbf{K}$ corresponding to one observation k are summarized in the column vector k of $\mathbf{K}$. The change in the entries of this vector due to the increase in the error of observation k in the iteration process is obtained by subtracting the vector after the error modification from the vector before that. Negative difference values indicate that the influence of the observation k on a particular state vector entry j is reduced, while a positive difference would indicate an increase of influence. If no change in the gain matrix elements is found, the change in the error for observation k has no effect on the state vector. This scheme of comparing the effects of changes in the measurement error model on the gain matrix is adopted in the following.

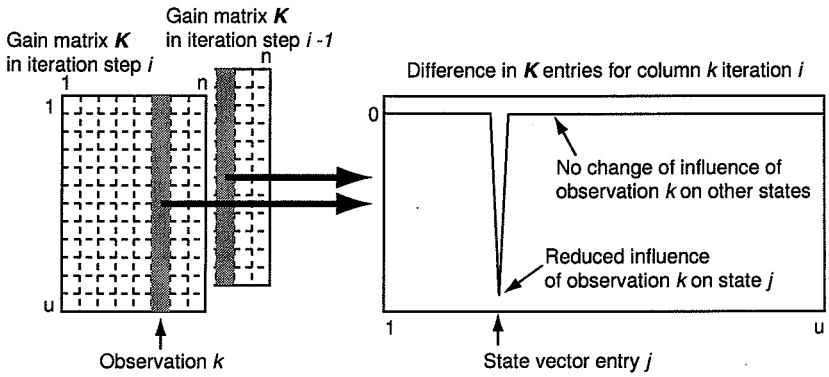


Figure 6.14: *Changes in gain matrix caused by increased measurement variance in one iteration step.*

Both examples presented in the following refer to the moving network point P10 in epoch 2. In this epoch P10 is measured from the two stations P3 and P4 with one spatial distance measurement, one horizontal and one vertical angle measurement each. The point is badly determined in this epoch 2 as has already been seen in Figure 6.3 and Figure 6.5.

Gross error in poorly controlled observation DMES P3–P10
A gross error of 1.5 mm is deliberately introduced to the poorly controlled spatial distance observation between P3 and P10 in epoch 2.

In the process of stochastic stabilization the gross error should be identified and its influence on the parameter estimation reduced. The method described in Chapter 5.5.3 analyzes the observation residuals. As the gross error has been added to a poorly controlled observation (redundancy number is 0.39), the error is likely to affect also other observation residuals. The localization step for the gross error detection analyzes the squared residuals which are tested against the corresponding quantile of the χ^2–distribution. For the example at hand up to four possibly outlying observations can be identified. Not only the truly erroneous spatial distance observation P3–P10, but also the spatial distance P4–P10 and both horizontal angle measurements P3–P10 and P4–P10. The number of suspected erroneous observations depends on the choice on how to adapt the level of significance in the derivation of the statis-

Table 6.6: *Small simulated network – Moving network point P10 – epoch 2: Gross error in observation 5 DMES P3–P10. Statistical test terms and redundancy numbers for observation 5 (DMES P3–P10), 9 (DMES P4–P10), 48 (ANGL P3–P10) and 52 (ANGL P4–P10).*

Observation	$\Omega^2_{\mathbf{v}_{GM,j}}$	$r_{\mathbf{z},i,k}$	$\chi^2_{1-\bar{\alpha},h_j}$ with $\bar{\alpha} = \alpha$	$\bar{\alpha} = \frac{\alpha}{n_i}$	$\bar{\alpha} = \frac{\alpha}{R_{\mathbf{z},i}}$	$\bar{\alpha} = \alpha \cdot (1 - r_{\mathbf{z},i,k})$
5	17.06	0.39	3.84	13.40	11.41	4.67
9	9.64	0.28	3.84	13.40	11.41	4.39
48	10.81	0.92	3.84	13.40	11.41	8.29
52	13.59	0.78	3.84	13.40	11.41	6.49

tical terms. As discussed in the previous Chapter 5.5.2 it is advisable to adapt the level of significance in a process applying repeated individual tests. The statistical test terms derived in the localization step are used to inflate the variance values of suspected erroneous observations (and system noise errors), i.e. in the stochastic stabilization process as defined in (5.109), (5.110) and (5.111). The variance inflation factor T_j is derived from the ratio of the squared residual term and the corresponding χ^2–quantile. A sensible choice for the adaptation of the level of significance for the statistical tests is vital to the adaptive KF employing stochastic stabilization.

The effect of the discussed choices to adapt the level of significance in the statistical tests can be seen for the example at hand in Table 6.6.

- $\bar{\alpha} = \alpha$: If the level of significance $1 - \alpha$ is not adapted at all, the χ^2 – test thresholds are smallest. All four observations would be suspected to contain a gross error and large correction factors $T_j = \frac{\Omega^2_{\mathbf{v}_{GM,j}}}{\chi^2_{1-\bar{\alpha},h_j}}$ would apply.
- $\bar{\alpha} = \frac{\alpha}{n_i}$ and $\bar{\alpha} = \frac{\alpha}{R_{\mathbf{z},i}}$: As n_i and $R_{\mathbf{z},i}$ are generally large numbers the values for $\bar{\alpha}$ become very small and the test thresholds increase. Applied to the example two observations would be considered to contain a gross error but the variance correction factors would be very similar and no differentiation would be made between the truly erroneous observation and the wrongly suspected observation.

Table 6.7: *Small simulated network – Moving network point P10 – epoch 2: Gross error in observation 5 DMES P3–P10. Statistical test terms and changes in observation error information due to stochastic stabilization for observation 5 (DMES P3–P10), 9 (DMES P4–P10), 48 (ANGL P3–P10) and 52 (ANGL P4–P10).*

Observation	$\Omega^2_{\mathbf{v}_{GM,j}}$	$\chi^2_{1-\bar{\alpha},h_j}$	T_j	r_k old	r_k new
5	17.06	4.67	3.65	0.2 mm	7.7 mm
9	9.64	4.39	2.19	0.2 mm	1.8 mm
48	10.81	8.29	1.31	7^{cc}	26^{cc}
52	13.59	6.49	2.10	7^{cc}	57^{cc}

- $\bar{\alpha} = \alpha \cdot (1 - r_{\mathbf{z},i,k})$: If the α is scaled with the factor $(1 - r_{\mathbf{z},i,k})$ as suggested in (5.108) the effect on other observations can be reduced. As the redundancy number is smaller for the poorly controlled observation, the corresponding value for $\bar{\alpha}$ is larger than for better controlled observations. This causes the test to be more sensitive in the case of poorly controlled observations.

The resulting test terms based on the adapted level of significance $1 - \bar{\alpha} = 1 - \alpha \cdot (1 - r_{\mathbf{z},i,k})$ are given in Table 6.7. The test terms T_j are used to inflate the measurement variances as defined in (5.110). The effect from this change in the stochastic information of the observation containing the gross error on the gain matrix entries is illustrated as outlined before, in Figure 6.15. In this illustration in can be seen that the effect is most significant for the entries in the state vector corresponding to P10, namely the position and velocity entries for the x-coordinate. The influence of the observation on these state vector components has been reduced. After the gross error detection and the stochastic stabilization process the redundancy number of the erroneous observation is 0.99 which implies that the observation has no influence in the estimation process.

Gross error in ANGL P3–P10 In a second small example an error of 60^{cc} is deliberately introduced to the horizontal angle measurement P3–P10. This observation is much better controlled than the observation in the example before. In the gross error detection only this single observation is suspected as containing a gross error. Due to the good

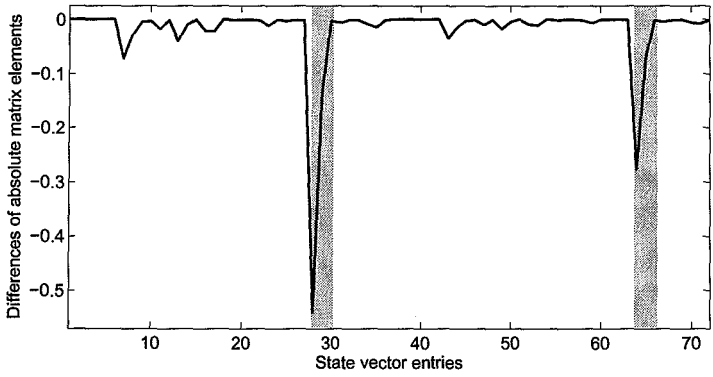


Figure 6.15: *Small simulated network – Moving network point P10 – epoch 2: Changes in gain matrix caused by increased measurement variance – one iteration. Illustration of column of gain matrix corresponding to observation dMES P3 – P10. The state vector entries corresponding to the position and velocity parameters respectively for P10 are indicated by the grey areas.*

Table 6.8: *Small simulated network – Moving network point P10 – epoch 2: Gross error in observation 48 ANGL P3–P10. Statistical test terms and changes in error information due to stochastic stabilization for observation 5 (DMES P3–P10), 9 (DMES P4–P10), 48 (ANGL P3–P10) and 52 (ANGL P4–P10).*

Observation	$\Omega^2_{\mathbf{v}_{GM},j}$	$\chi^2_{1-\bar{\alpha},h_j}$	T_j	r_k old	r_k new
5	0.87	4.67	0.19	0.2 mm	0.2 mm
9	1.67	4.39	0.38	0.2 mm	0.2 mm
48	26.58	8.29	3.21	7^{cc}	173^{cc}
52	5.17	6.49	0.80	7^{cc}	7^{cc}

control in this observation the error is clearly detected and does not affect other observation residuals as can be seen in Table 6.8 listing the results for the statistical terms using the adapted level of significance $1-\bar{\alpha} = 1-\alpha\cdot(1-r_{\mathbf{z},i,k})$.

Stochastic stabilization is achieved by changing the single observation error. The effect of this change on the gain matrix is visualized in Figure 6.16. The entries corresponding to the position and velocity

states of the x- and y-coordinate of P10 are most affected by the change of the measurement error.

The redundancy number of the erroneous observation is 1.0 after the stochastic stabilization process.

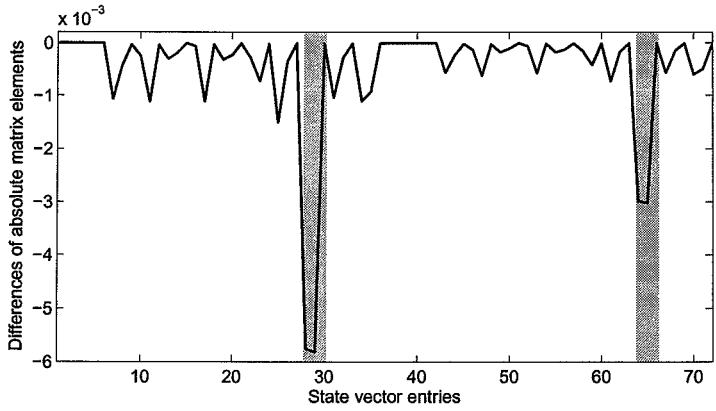


Figure 6.16: *Small simulated network – Moving network point P10 – epoch 2: Changes in gain matrix caused by increased measurement variance – one iteration. Illustration of column of gain matrix corresponding to observation ANGL P3 – P10. The state vector entries corresponding to the position and velocity parameters respectively for P10 are indicated by the grey areas.*

These two examples have shown the effect of gross errors in the observation data on the KF gain matrix if stochastic stabilization is carried out. The influence of such errors on the KF estimation result is decreased.

System noise error adaptation

Adapting the system noise error in the stochastic stabilization process affects the gain matrix $\mathbf{K}$ as seen in Chapter 5.5.5. Using the example of the moving network point P10 in the original (i.e. unchanged) data set, this effect is illustrated in the following. The movement of P10 in z-direction in the last measurement epoch 9 is larger than expected by the system description which could already be seen in Figure 6.4 and

Figure 6.9. The same applies to the other moving network points P3, P4 and P9. The system description for all four points has to adapt to the unexpected change.

Discrepancies are detected by the global model test which can be localized in the predicted state vector residuals $\mathbf{v}_{\mathbf{x},i}$ of the z-coordinates of all four moving network points P3, P4, P9 and P10. In several iteration steps the corresponding system noise variances are inflated. After each change the KF adjustment for this data epoch is repeated, affecting also the gain matrix. If the global model test still indicates a discrepancy between the system description and the measurement data the stochastic stabilization is repeated. This process is iterated until the global model test is passed.

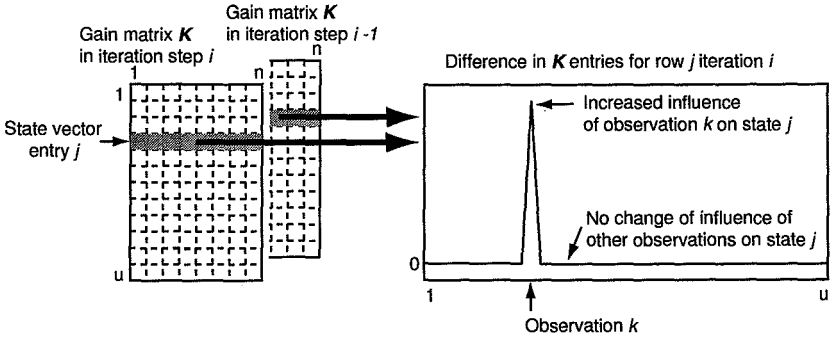


Figure 6.17: *Changes in gain matrix caused by increased measurement variance in one iteration step.*

For P10 five iteration steps are necessary to adapt the system description sufficiently. The derivation of the statistical test terms is based on the adapted level of significance $1 - \bar{\alpha} = 1 - \alpha \cdot (1 - r_{\hat{\mathbf{x}}_{i|i-1},k})$. The redundancy numbers for the predicted state vector elements $r_{\hat{\mathbf{x}}_{i|i-1},k}$ are defined in (5.106). A small predicted state redundancy number indicates an already well known system description. By scaling α by $(1-r_{\hat{\mathbf{x}}_{i|i-1},k})$, $\bar{\alpha}$ is larger for well known system states than for other states which are not so well known yet. The corresponding test threshold is smaller for well known system states which makes the test decision in these cases more sensitive, i.e. these system states are more likely suspected to be the cause of a misspecification in the system description than others.

This can avoid the effect that the KF learns the system description too fast too well which is a common problem in many KF applications, [15, 35].

In Table 6.9 the statistical terms are given for the five necessary iteration steps to adapt the system description for the z-coordinate of P10.

Table 6.9: *Small simulated network – Moving network point P10 – epoch 2: System noise variance adaptation z-coordinate of P10. Statistical test terms and changes in observation error information.*

Iteration	$\Omega^2_{\mathbf{v}_{GM,\,j}}$	$\chi^2_{1-\overline{\alpha},h_j}$	T_j	$q_{i-1,j}$ old [mm/dt^3]	$q_{i-1,j}$ new [mm/dt^3]
0	10.33	6.58	1.57	0.005	0.011
1	9.96	6.59	1.51	0.011	0.023
2	9.58	6.60	1.45	0.023	0.048
3	8.98	6.64	1.35	0.048	0.095
4	8.10	6.76	1.20	0.095	0.172
5	7.01	6.97	1.01	0.172	0.285

In the first localization (iteration 0) the quadratic term for the predicted state vector residual $\Omega^2_{\mathbf{v}_{GM,\,P10_z}}$ exceeds $\chi^2_{1-\overline{\alpha},h_{P10_z}}$ and the ratio test term T_{P10_z} is larger than one. The increase in the system noise variance is derived as in (5.111), resulting in an increase from 0.005^2 to 0.011^2 and eventually, after five iterations to 0.285^2 when the global model test passes. In the next epoch of measurement data the system noise error is reset to its initial value. If the global model test for this new measurement epoch indicates any discrepancies the iteration to achieve stochastic stabilization starts again from these initial system noise error values.

Figure 6.17 illustrates how the effect of the change in the system noise error on the gain matrix can be visualized, similar to Figure 6.14. The entries in the gain matrix $\mathbf{K}$ corresponding to the state parameter j are summarized in the row vector j of $\mathbf{K}$. The change in the entries of this vector due to the increase of the system noise error for state j is obtained by subtracting the vector after the error modification from the vector before.

If the system noise error is increased, the uncertainty in the system description increases and more influence is given to the observations. Positive difference values indicate that the influence of an observation

k on the state vector entry j is increased. If there is no change in the gain matrix elements, the observations maintain the same influence as

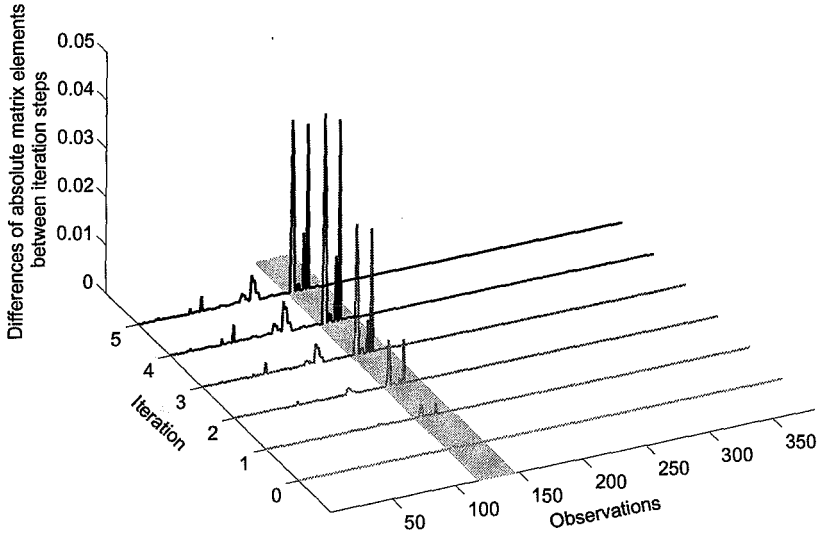


Figure 6.18: *Small simulated network – Moving network point P10 – epoch 9: Changes in gain matrix caused by increased system noise variance – five iteration steps. Illustration of gain matrix row vector corresponding to z-coordinate of point P10. The height difference observations are indicated by the grey area.*

Figure 6.18 shows the differences in the row vector elements corresponding to the position state of the z-coordinate of P10 for all five iteration steps. The effect of the gain matrix entries in the first iteration steps is small. The change of the system description induced by the inflation of the system noise variance is insufficient to overcome the discrepancy between the system description and the measurements. With continued iteration the system noise variance is increased, provoking a larger influence on the gain matrix entries until the system description has sufficiently adapted and the global model test passes.

The most significant differences considering the observations can be seen in the range of the height difference observations, indicated by a

grey area in Figure 6.18. Especially three individual observations (index 123, 124 and 135) stand out: These observations can be identified as the three height difference measurements connecting P10 to other network points. As expected it can be concluded that due to the increase in the uncertainty in the system description of the z-coordinate of P10, more weight is given in the KF to observations which determine this unknown.

In the three examples given here the effect of stochastic stabilization on the Kalman gain matrix in the adaptive KF implementation have been illustrated. Changes in the stochastic models of the measurements and the system noise affect the gain matrix. By adapting the level of significance in the related statistical tests employing the redundancy numbers of the observations and the predicted system states, possibly negative effects on the KF performance of poorly controlled observations and too well known system state parameters can be reduced.

6.2.4 Including additional information in the adaptive KF

As discussed in Chapter 5.6 additional information can be included in the KF adjustment. This information can be used to include empirical knowledge about the expected point deformations. In this example the four moving floor points are expected to move exclusively in vertical direction. The results for network point P10 are compared to the results of the adaptive KF without this additional information. The KF results again consider only the simpler system model SMI.

In Figure 6.19 and Figure 6.20 individual coordinate components are compared to the true point positions over all ten measurement epochs. The corresponding $1\,\sigma$ error information is plotted by dashed lines. In Figure 6.19 the results for the case of no additional information can be seen. They are equivalent to the KF results for the SMI presented in Figure 6.3, Figure 6.4 and Table 6.3. Again it can be seen that in epoch 2 the point is badly determined in x-direction.

In Figure 6.20 the additional information that the floor points are expected to move only in vertical direction is included in the KF adjustment. The differences to the true values are smaller which is also reflected in the smaller corresponding errors. The effect of bad configuration in epoch 2 is reduced.

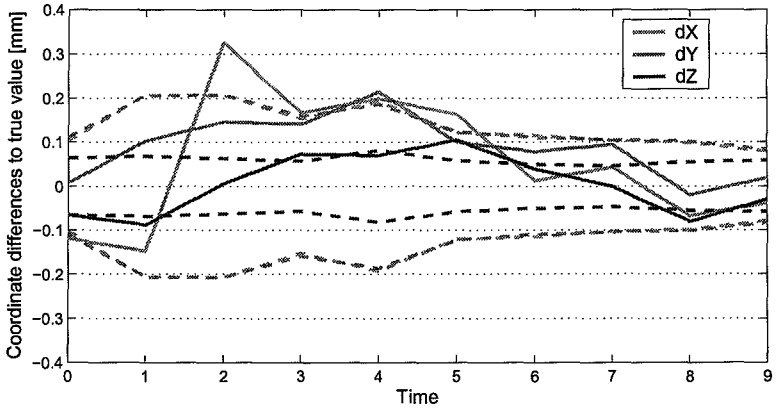


Figure 6.19: *Small simulated network: Comparison of KF results (SMI) to true positions – no additional information: Network point P10, individual coordinate differences and corresponding 1σ error information (dashed lines).*

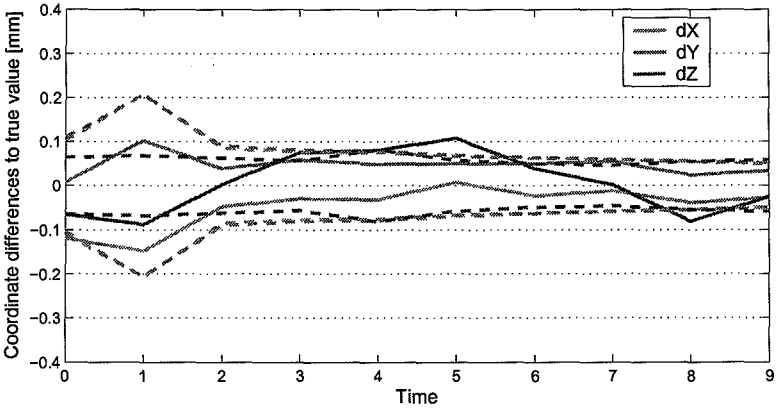


Figure 6.20: *Small simulated network: Comparison of KF results (SMI) to true positions – additional information included: Movement only in vertical direction. Network point P10, individual coordinate differences and corresponding 1σ error information (dashed lines).*

The solution of the KF can be supported by including additional information which is especially effective in situations of bad measurement configuration or when no measurement information is available (here in epoch 4).

6.2.5 Conclusion for the small simulated network

The results for the very simple example presented in this section show the performance of the adaptive KF implementation for survey data analysis in a deforming network. Changing network measurement configurations are easily handled in the KF implementation and point deformations can be derived as they are part of the system description of each point.

The comparison of KF and individual epoch standard least squares (SE LSQ) results with respect to true positions shows that the KF results are more accurate and precise. KF estimates are less affected by bad network configurations, poorly controlled or uncontrolled observations and missing observations. KF estimates are available for points that are included in the system description also if no measurements are available in an epoch, by means of the predicted values. The corresponding, predicted error estimation gives useful information about the point's accuracy for the planning of survey measurements.

The reliability in the network observations is increased on average by 10% for the application of the KF compared to the SE LSQ results. Bad measurement configurations reduce generally the reliability of a network and its observations. Any gross errors in these observations would be difficult to detect. By applying the adaptive KF in kinematic setup this situation can be considerably improved by maintaining the reliability once a good system model description is established.

The effect of stochastic stabilization in the adaptive KF has been discussed by illustrating its effect on the KF gain matrix. It has been shown that the adaptation of the level of significance in statistical tests using the redundancy numbers of the observations and the predicted system states helps to reduce negative effects of poorly controlled observations and too well known system state parameters. The stochastic stabilization process is the key property of the adaptive KF introduced in this thesis avoiding possible filter divergence and ensuring correct identification of the kinematic system model representing a deforming geodetic network.

It has been shown that the results of the adaptive KF can be further improved in terms of accuracy and precision if additional information about the point movements is included.

6.3 Simulated ATLAS cavern network data

A simulation of the ATLAS cavern network is carried out and the data are analyzed to test the adaptive KF algorithm on a more complex and larger data set, similar to the expected real situation in the ATLAS cavern. The design layout of the metrological ATLAS cavern network, information about expected cavern and access structure deformation and the projected installation schedule are used for the simulation of survey data over a period of 30 months. This time period corresponds to the originally envisaged installation time for the ATLAS detector in the cavern.

15 epochs of data are simulated each in different network and measurement configurations depending on the assumed installation schedule and the simulated network deformation. Deformations of network points are derived from a simplified deformation model described in the following section. The deformations are translated to changing network point coordinates which are thereafter used for the simulation of corresponding survey data. It is emphasized here that the simulation of deformations is dedicated to the purpose of generating survey data representing possible scenarios. It is not to be understood as a simulation of the real structural cavern deformation.

The simulated survey data is analyzed to identify critical network parts and properties. Subsequently the data is processed with the adaptive KF algorithm implementation. Results are presented for selected features mainly by plots.

6.3.1 Network data simulation

Several tools to simulate the ATLAS cavern network configuration are developed in MATLAB©6.5. These applications are based on information about the design network, the detector installation process and expected deformation of the cavern and access structures. Data input files defining the changing network and measurement configurations are created in the standard LGC [31] format. The network deformation

is represented by changing network point coordinates. The input files are used to actually simulate survey data for the given configuration in the LGC software under the option **SIMU**. The simulated data is subsequently adjusted and analyzed in the adaptive KF MATLAB application described in Section 6.1.

Network configuration

The design of the metrological ATLAS cavern network as described in Chapter 2.4, the installation schedule summarized in Chapter 2.3 and approximate envelope descriptions of obstacles are used to determine different network configurations over the planned installation period. Obstacles for sightings between points include cavern and other underground structure walls, access structure gangways and detector parts. Limitations on the sightings are determined by spatial intersection with these possibly obstructing objects.

An example of simulated sighting limitations is shown in Figure 6.21, which illustrates possible observations for epoch 6 (time 12 – corresponding to installation phase 7, see also Figure 2.14).

Network deformation

The deformation of the network points is simulated by establishing a simplified, combined model of cavern and access structure deformation. The derived deformation vectors are translated to changing coordinates for the network points.

Cavern deformation A simple deformation model for the cavern structure is derived from information discussed in Chapter 2. Although the civil engineering calculation results and designs of the final cavern linings stated that no deformations would affect the cavern lining itself, small deformations of the lateral cavern walls are expected and considered in the simulation with an empirical value observed in former LEP experiments (cavern shrinkage by 1 mm/year, [29]). Additionally a sagging of the cavern floor by 3.6 mm/year for the first 24 months of installation is expected. The simplified cavern floor deformation is modelled by a quadratic surface, with its maximum at the center of the cavern, decreasing towards the cavern walls. After the first 24 months

It has been measured twice in horizontal position in early epochs and then has not been measured for almost two years. In Figure 6.44 it can be seen that a small movement in the x-and y-direction in epoch 3 (time 14) causes the KF to project this movement over the long time period without measurement data. The KF update with measurements in epoch 42 (time 37.8) corrects this wrongly identified deformation but the error estimates remain at a high level especially in the y-direction caused by the large predicted uncertainty for the missing data epochs. This situation is similar to the previous case but more extreme here as the point movement estimation is based on only two measurement epochs and the data gap is even longer.

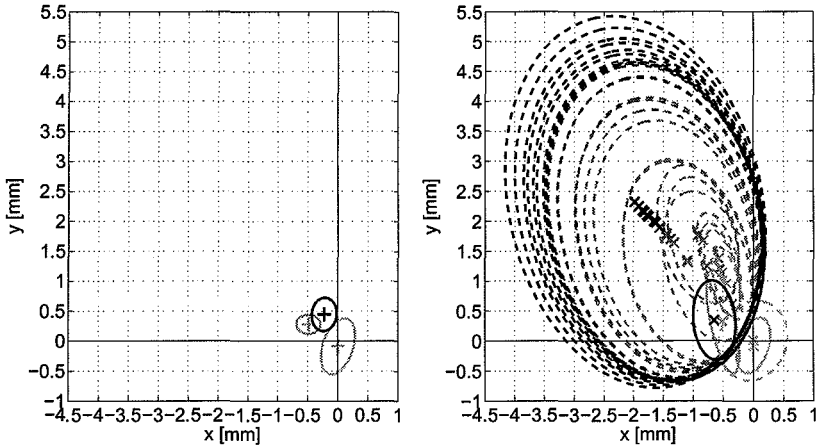


Figure 6.44: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for the cavern floor point A11+5 in the xy-plane.*

In the vertical positions presented in Figure 6.45 the SE LSQ results indicate a movement in z-direction, first positive in z-direction and then reversing its direction to resume approximately the same height level (+0.3 mm). The +z movement of approximately 1 mm is observed by geometrical levelling measurements and the KF follows this positive trend only very slowly. In epoch 19 the SE LSQ result is similarly off as in the previous case of point C22-15, caused by wrong approximate coordinates for the freely adjusted levelling network.

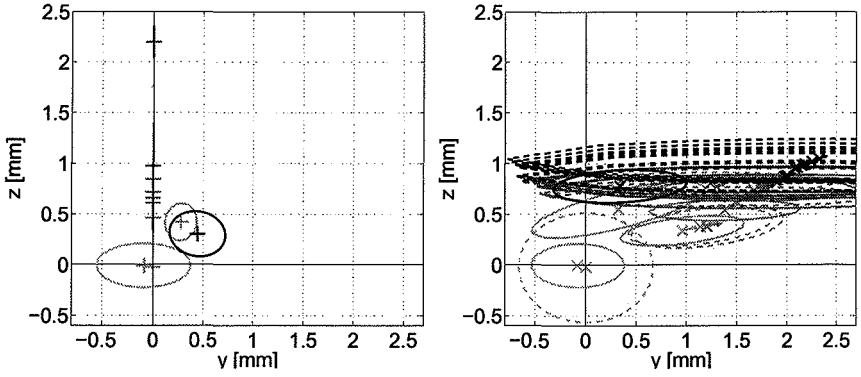


Figure 6.45: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for the cavern floor point A11+5 in the yz-plane.*

The last estimate of the KF indicates a +z movement for this point of 0.8 mm. The KF only slightly adapts the estimate from the last predicted value of 1 mm. This point has been determined in this last epoch from two temporary stations and one network point by angle and spatial distance measurements (no levelling). The measurements are not sufficiently redundant to force the KF result to the presumably correct position indicated by the SE LSQ result. The KF estimates up to epoch 19 correspond to results seen in a detailed independent analysis of the levelling data of the ATLAS cavern, [14].

The KF redundancy numbers for the position state estimates plotted in Figure 6.46 are similar to the case of the point C22-15 before. As here only three epochs of horizontal information are available, the redundancy numbers for the horizontal states always resume a value close to one when new measurements are available. No trust is given to the system description for the point's horizontal position. Also here the vertical position is better determined due to the extra levelling measurements, resulting in an increasingly reliable vertical point position dscription.

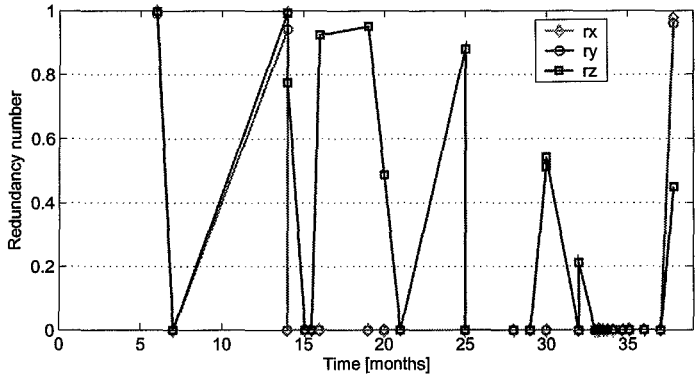


Figure 6.46: *Real ATLAS cavern network data: KF redundancy numbers of position state entries for the cavern floor point A11+5.*

HO access structure point HOC06	
Measured in	5 epochs
3D	5 epochs (time 7, 14, 21, 28 and 37.8)
levelling only	none

This point is mounted on the access structure facing the end side C. Its horizontal positions are illustrated in Figure 6.47. A significant motion in +y-direction can be seen which is identified by the KF, but also could have been derived from SE LSQ results.

A significant motion in -z-direction is visible in Figure 6.48 which is estimated in epoch 10. The large movement propagates into the subsequent epochs 11 and 12 with no measurement data available. The estimation is corrected by the next KF update in epoch 13 when new measurement data is available.

Considering the position state KF redundancy numbers shown in Figure 6.49 it can be seen that as this point is always observed in 3D, no large differences exist between the redundancy numbers for individual coordinate components. In the very last epoch the redundancy number for the x-position remains on a very low level indicating that the KF does not consider the measurement information for this coordinate component very reliable (i.e. the observation is down-weighted in the iterative gross–error detection) and puts more weight on the system prediction for this value. This also explains a slightly inflated error ellipse

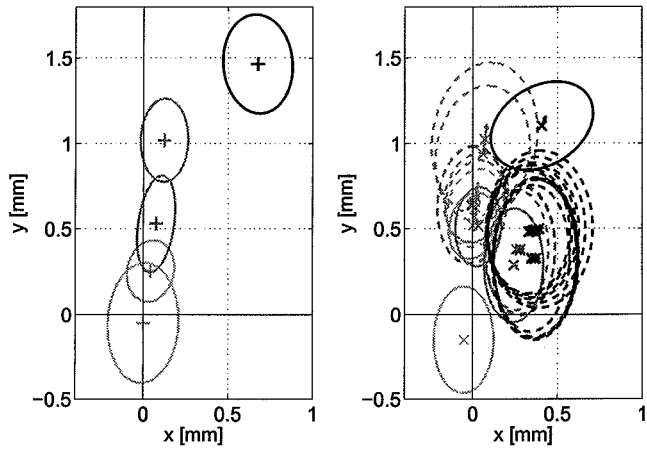


Figure 6.47: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for access structure point HOC06 in the xy-plane.*

in Figure 6.47 for this positions as well as the difference in x-direction between the SE LSQ and KF estimates.

Access structure points show generally very large movements, not always in such a systematic way as for point HOC06. More random movements can be observed which are often caused by unexpected forces acting on these metal structures due to the very active installation work in close vicinity of these structures. If the point movements are more random and no systematic long term movement is present, it is more difficult for the KF to estimate movements, as will be seen in the next example.

Coordinates of points on access structures in general have to be considered of inferior quality compared to cavern wall points, as point movements are not easily predicted. Using these points as reference points for survey measurements in the cavern should always be verified and controlled by making the connection to more reliable network information.

Concerning the effect of fixing and constraining the access structures to the cavern walls, no conclusion can be drawn from the results for these point. The point movements are too erratic to derive any systematic deformation of these structures.

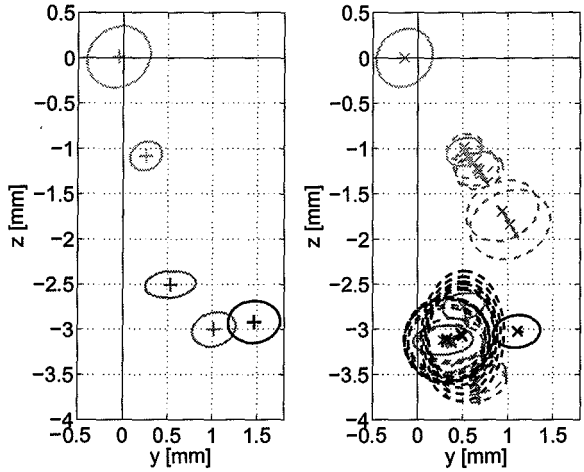


Figure 6.48: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for access structure point HOC06 in the yz-plane.*

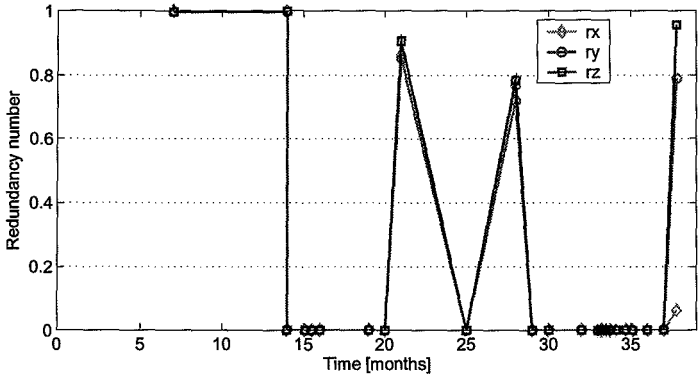


Figure 6.49: *Real ATLAS cavern network data: KF redundancy numbers of position state entries for the access structure point HOC06.*

HO access structure point HOA16	
Measured in	6 epochs
3D	5 epochs (time 7, 14, 21, 28 and 37.8)
levelling only	1 epochs (time 14)

In Figure 6.50 and Figure 6.52 a more random point movement in the horizontal and vertical directions for point HOA16 can be observed. The SE LSQ results indicate a significant movement in -x-direction which is only slowly followed by the KF. In this point's first measurement epoch in epoch 3 (time 7) it is sighted only from stations on the other far end of the cavern resulting in small angles of intersecting rays and thus low precision in the y-direction of the position estimates. At this stage the HO structures had already been installed but gangways for access of the lateral cavern walls or HS structures were not yet in place prohibiting the installation of any network points in these areas. The point movement in the following epochs in the horizontal position are not easy to identify for the KF which is also due to the poor estimation accuracy in the first epoch.

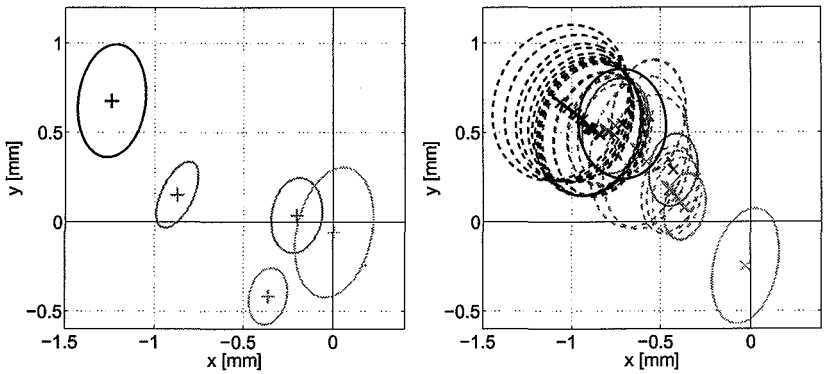


Figure 6.50: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for access structure point HOA16 in the xy-plane.*

The situation is similar in the determination of the vertical positions shown in Figure 6.52. The point movement in the -z-direction is very erratic and cannot be followed very well by the KF which results in a large difference between the SE LSQ and the KF result in the last epoch.

Figure 6.51 illustrates the KF redundancy numbers for the position state entries for this point which follow a similar pattern as as for point HOC06.

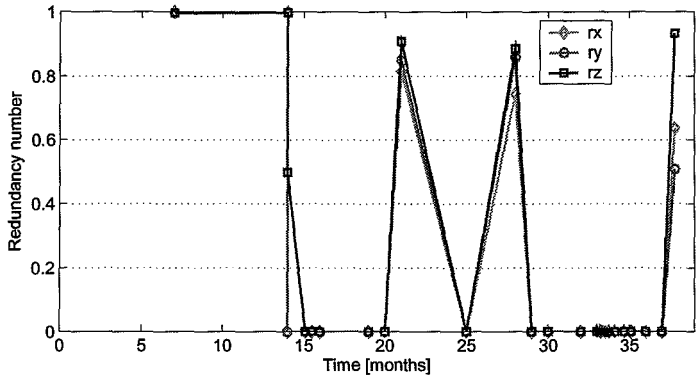


Figure 6.51: *Real ATLAS cavern network data: KF redundancy numbers of position state entries for the access structure point HOA16.*

Survey gallery point POTTUNUPSC	
Measured in	10 epochs
3D	5 epochs (time 6, 7, 14, 21 and 37.8)
levelling only	5 epochs (time 14, 16, 25, 30, 32)

This point is located at the far end of the survey gallery on side C. It is poorly determined in the horizontal positions in epoch 2 (time 7), as can be seen in Figure 6.53. In this case the point is only sighted from one station in the cavern. No connection to LHC tunnel reference points is available in this epoch. The KF result is only slightly affected, as the KF includes information for this point from one previous epoch when this point has been well determined. For this epoch the result is more accurate and precise compared to the SE LSQ result.

A small horizontal movement in the +x-direction is identified for this point which is estimated first too large by the KF, causing the predicted positions in the subsequent epochs to drift off. The situation is corrected in the update in epochs 10 (time 21) and 42 (time 37.8) when measurement data is available for this point.

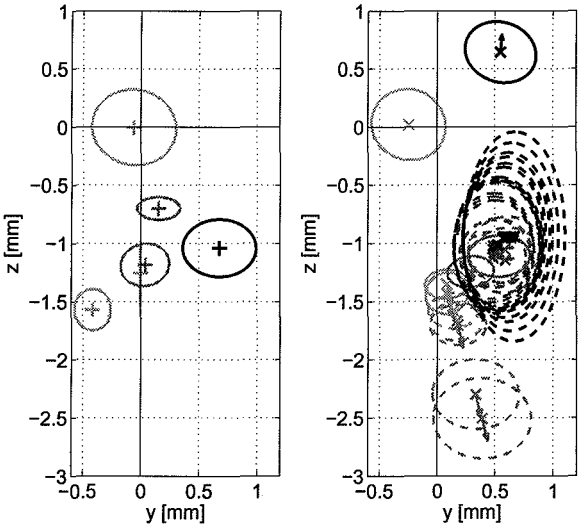


Figure 6.52: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for access structure point HOA16 in the yz-plane.*

Similarly to the horizontal positions the vertical positions show in Figure 6.54 that the point is also determined poorly in the vertical direction in epoch 2. Overall no significant vertical movement can be identified for this point.

Figure 6.55 shows an effect of the bad measurement configuration in epoch 2 (time 7) on this point's position state KF redundancy numbers. The redundancy is already reduced in the second epoch which reflects little confidence in the measurements. In the subsequent epochs of available measurement data the reliability in the state estimates is increased as the redundancy numbers decrease. The long break between epoch 10 (time 21) and epoch 42 (time 37.8) reduces the reliability in the system description for the horizontal positions as seen above for the position results.

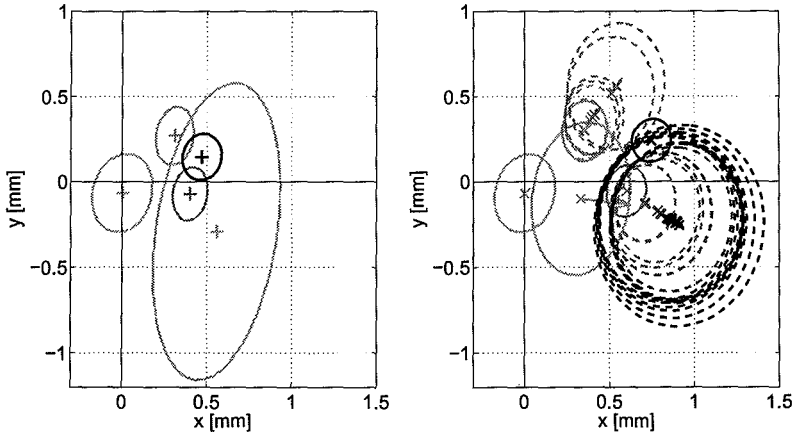


Figure 6.53: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for survey gallery point POTTUNUPSC in the xy-plane.*

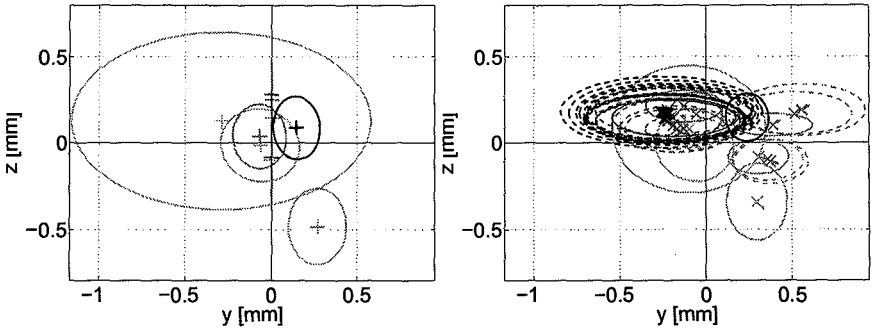


Figure 6.54: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for survey gallery point POTTUNUPSC in the yz-plane.*

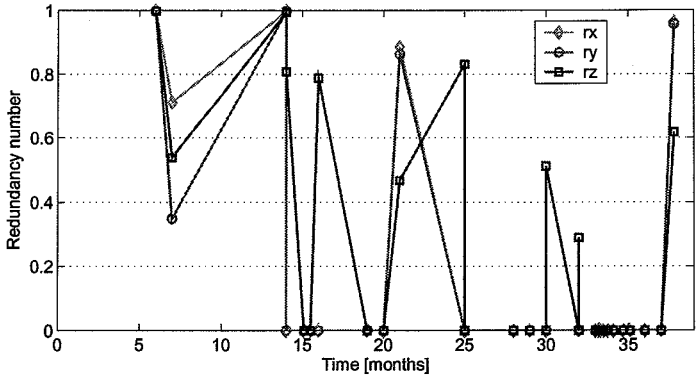


Figure 6.55: *Real ATLAS cavern network data: KF redundancy numbers of position state entries for the survey gallery point POTTUNUPSC.*

USA lateral wall points at beam level: MA±xx-00

This group of points is mounted on the lateral cavern wall on side USA, arranged in one line and approximately at the same height. The point names indicate their position in x- and z-direction with respect to the IP (origin of the local coordinate system), e.g. MA+21-00 is approximately 21m from the IP in x-direction and at approximately the same height as the IP.

The results for this group of points are interesting as they show the deformation of the cavern wall. Results for two individual points are presented in detail.

Lateral wall point MA+21-00	
Measured in	10 epochs
3D	5 epochs (time 14, 15.5, 21, 28, 37.8)
levelling only	5 epochs (time 14, 16, 25, 30, 32)

A large movement in +x-direction is estimated in epoch 7 and 10 for the cavern wall point MA+21-00 as can be seen in Figure 6.56. This movement is propagated to the following epochs and is not sufficiently corrected in epoch 14, but in epoch 42 (time 37.8). The resulting de-

formation vector for this last epoch is very small in reaction to the necessary correction. The corresponding KF redundancy numbers for the state estimates in Figure 6.58 show that the reliability in the x-position is smaller (represented by large redundancy numbers in the measurement update epochs) than for the other position components. This reflects the adaptation of the KF to the large deformation in x-direction by increasing the corresponding system noise error.

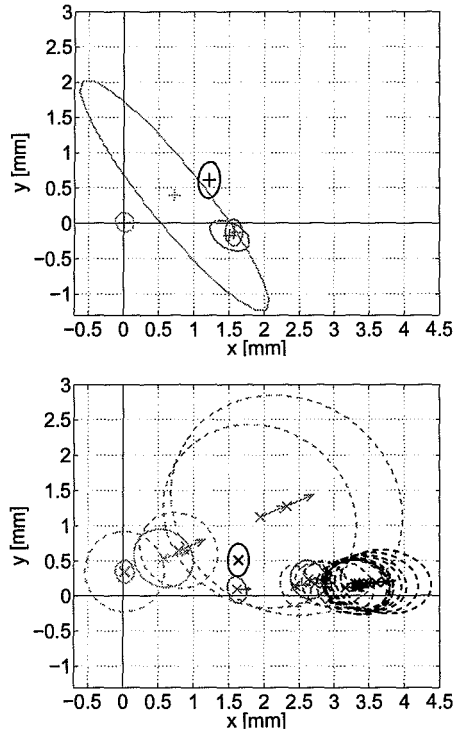


Figure 6.56: *Real ATLAS cavern network data: Comparison of SE LSQ results (top) and KF results (bottom) for cavern wall point MA+21-00 in the xy-plane.*

A significant vertical movement in the -z-direction can be seen in both the SE LSQ and the KF results in Figure 6.57.

A large uncertainty in the position for this point arises from a measurement in a small survey measurement (i.e. not part of a dedicated network measurement campaign) in epoch 6 (time 15.5).

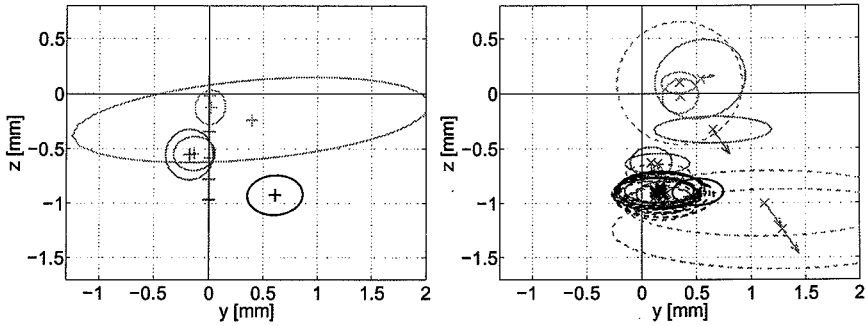


Figure 6.57: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for cavern wall point MA+21-00 in the yz-plane.*

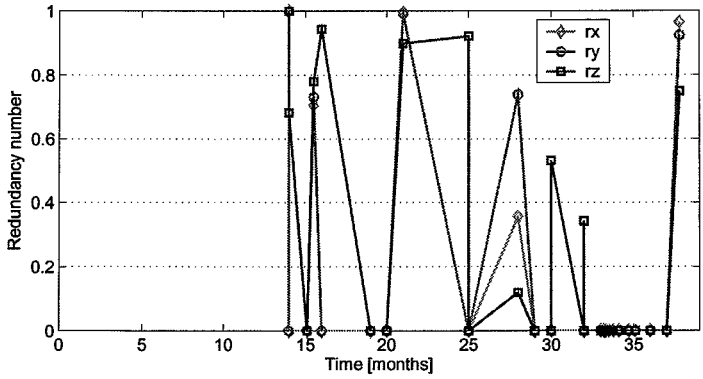


Figure 6.58: *Real ATLAS cavern network data: KF redundancy numbers of position state entries for the cavern wall point MA+21-00.*

Lateral wall point MA+06-00	
Measured in	8 epochs
3D	4 epochs (time 14, 15.5, 21, 37.8)
levelling only	4 epochs (time 14, 16, 25, 32)

In the horizontal positions depicted in Figure 6.59 a large movement in +y-direction is visible. It is distinctively identified in epoch 10 (time 21). The KF propagates this movement to the subsequent epochs until the next update is available in epoch 42 (time 37.8). In this epoch the estimate is corrected and the resulting velocity vector estimate is small due to the previously predicted large point movement.

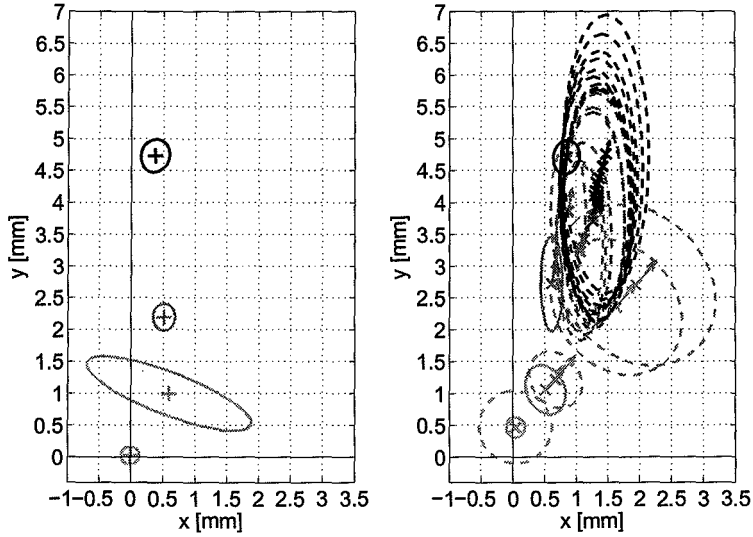


Figure 6.59: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for cavern wall point MA+06-00 in the xy-plane.*

A small but significant vertical movement in the -z-direction can be seen in Figure 6.60 in both the SE LSQ and the KF results.

The KF state redundancy numbers illustrated in Figure 6.61 for this point reflect the position results. The movement in y-direction and the corresponding elevated system noise component for this coordinate re-

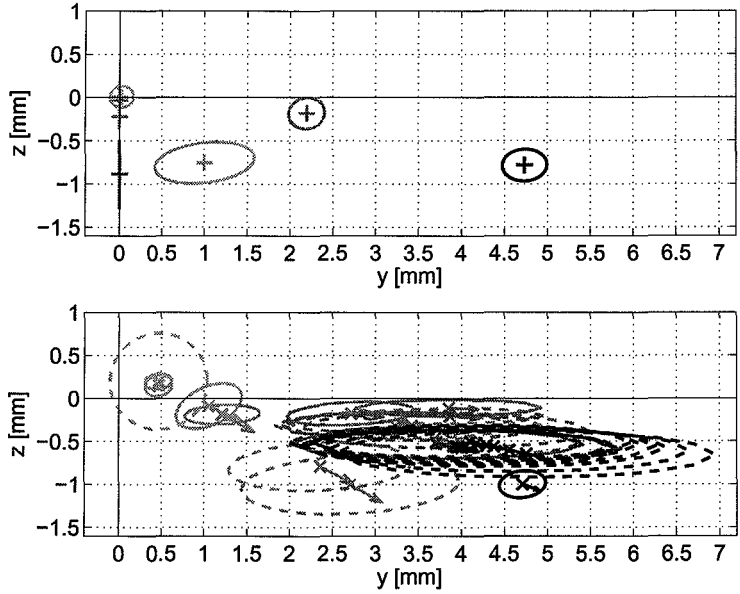


Figure 6.60: *Real ATLAS cavern network data: Comparison of SE LSQ results (left) and KF results (right) for cavern wall point MA+06-00 in the yz-plane.*

duce the reliability of the y-position. The reliability in the z-position is better, as more measurements for this coordinate component are available.

The total group of points along the USA side cavern wall shows a distinctive +y- and small -z-movement, as can be seen in Figure 6.62 showing the estimated point movements between epoch 4 and epoch 42 over a period of approximately 2 years.

The deformation vectors indicate a lateral deformation of the cavern wall, towards the center of the cavern, with its maximum at the center of the wall and smaller values close to the end walls.

The deformation vector for point MA+21-00 does not correspond very well with the overall deformation behavior both by direction and amount. The movement is dominantly in +x-direction and its amount is larger than the deformation of neighboring or symmetric points. It has to be considered that the platform for visitors to the ATLAS cav-

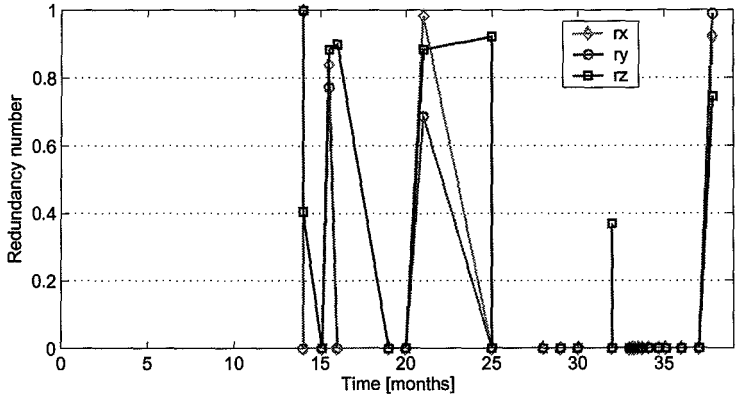


Figure 6.61: *Real ATLAS cavern network data: KF redundancy numbers of position state entries for the cavern wall point MA+06-00.*

ern is accessed from an entry between network point MA+21-00 and MA+18-00. The visitors' way to the platform passes directly by the foldable bracket for point MA+21-00. It can be assumed that the deformation of this point is mainly caused by the visitors passing even though the point is reasonably well protected and is not caused by cavern wall deformations.

Increase in reliability measured by comparison of observation redundancy numbers The overall mean observation redundancy numbers for the SE LSQ and the adaptive KF results can also be derived for the real ATLAS cavern data set. They are given in Table 6.17. In epochs with zero redundancy (all HLS measurements starting from epoch 20) for the SE LSQ solution the increase is in any case 100%. As the HLS measurement epochs only include the HLS measurements, the increase of the KF redundancy numbers over time represents the improvement of the system description in the HLS points which are only sparsely linked to the full network.

For the remaining epochs the increase is also largest for epochs with poorly controlled observations, e.g. small surveys in the cavern. This is not surprising as these observations are generally less precise and the configuration is worse than in dedicated network measurements. A

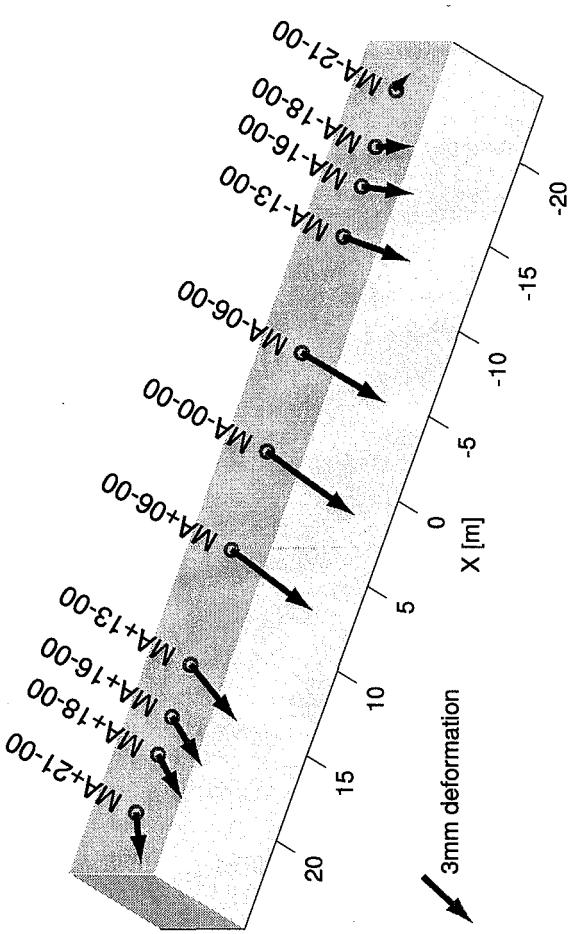


Figure 6.62: *Real ATLAS cavern network data: 3D point movements of lateral wall points MA±xx-00 estimated between epoch 4 (April 2004) and epoch 42 (March 2006).*

good system description in the adaptive KF augments the redundancy in these individual measurements and thus the reliability in the positions determined from these measurements. This shows the benefit of implementing the adaptive KF in kinematic setup for the deforming ATLAS cavern network. The reliability gains in dedicated network measurement campaigns are around 10% but for the small everyday measurements gains of at least 20% can be achieved.

Table 6.17: *Real ATLAS cavern network: Overall mean observation redundancy numbers for SE LSQ and KF solutions.*

Epoch	Mean redundancy SE LSQ	KF	Reliability gain in %	Epoch	Mean redundancy SE LSQ	KF	Reliability gain in %
0	0.86	0.86	0	22	0.00	0.60	100
1	0.72	0.72	0	23	0.00	0.70	100
2	0.77	0.80	4	24	0.00	0.76	100
3	0.85	0.85	0	25	0.00	0.80	100
4	0.44	0.50	11	26	0.00	0.83	100
5	0.43	0.54	21	27	0.00	0.84	100
6	0.37	0.42	11	28	0.00	0.86	100
7	0.45	0.46	4	29	0.00	0.87	100
8	0.50	0.51	2	30	0.00	0.88	100
9	0.54	0.59	9	31	0.00	0.88	100
10	0.81	0.84	4	32	0.00	0.89	100
11	0.51	0.53	4	33	0.43	0.94	54
12	0.42	0.43	2	34	0.30	0.49	39
13	0.66	0.71	8	35	0.00	0.24	100
14	0.31	0.49	36	36	0.54	0.63	14
15	0.54	0.55	3	37	0.00	0.01	100
16	0.53	0.56	5	38	0.00	0.11	100
17	0.87	0.91	4	39	0.00	0.28	100
18	0.42	0.48	13	40	0.00	0.25	100
19	0.48	0.55	13	41	0.00	0.33	100
20	0.00	0.04	100	42	0.74	0.79	6
21	0.00	0.34	100				

6.4.4 Conclusion for the real ATLAS cavern network

In this section the performance of the adaptive KF algorithm applied to real ATLAS cavern network data has been shown. It could be seen that the KF point position estimates are in general more accurate and

more precise than conventional SE LSQ solutions if sufficient data is available.

- In difficult measurement configurations the KF estimation is less affected than SE LSQ solutions as more information than just data of the single epoch is used to derive position estimates. Information from previous measurement epochs is represented in the system description that is propagated to following epochs.

- If no datum definition is included in the measurements the KF refers back to its system state description to relate the data to the datum definition. In SE LSQ a datum has to be defined in each epoch. In case the datum definition is done by choosing arbitrarily some network points, any error in these points' coordinates is directly propagated on the newly estimated point coordinates.

- Random and unsystematic point movements can be difficult to estimate for the KF especially if the measurement information provided to the update step is of poor quality.

- Low data sampling rates cause the KF estimations (predictions) to drift off. However, the error estimations represent the increased uncertainty in the position and point movement estimates.

The quality of this data set of approximately three years of survey measurements gathered in the environment of the ATLAS experimental cavern is much poorer than optimal, especially considering the horizontal positions. Still the KF gives good deformation estimations also in cases of few available data epochs. In very difficult situations, when factors like bad initial point determination, difficult measurement configurations and very low data sampling combine, the KF reacts slowly to unexpected movement changes. This can result in wrong position estimates. In these cases the SE LSQ results are more accurate.

It has been shown that also for this very complex and heterogenous data set a gain in reliability of approximately 10% can be achieved by applying the adaptive KF in kinematic setup. For epochs with less precise measurements in bad configurations this gain can be more than 20%.

6.5 Conclusions for the application and results of the adaptive KF

The application of the adaptive KF in kinematic setup to three different data sets has been presented in this chapter. The simulated data examples proved an enhanced performance of the adaptive KF implementation compared to individual epoch least-squares adjustment. The KF is not only more accurate and precise in estimating point positions, degrading or rank deficient network configurations have little effect on the point estimates. The reliability of the network expressed by redundancy numbers can be maintained in these difficult situations. It has been shown that the reliability gain for using the adaptive KF compared to SE LSQ is on average 10% and more if the measurement configurations are poor.

Deformation vectors for individual points are included in the system description and no additional analysis step is necessary to derive them. In epochs with no observation data available for a particular point, the prediction of the state vector and the corresponding error estimation give a powerful tool for the planning of network measurements and also for the quality management of the network.

The application of the adaptive KF to real ATLAS cavern network data for a period of more than three years showed that the algorithm can also be used on extremely difficult data sets. The benefits seen in the applications on the simulated data could also be observed in the real world situation. Real point movements that are not as systematic and clearly identifiable as simulated point movements could be detected and modelled. Nevertheless, low data rates in most of the network points compared to the simulated examples render the analysis of the data less conclusive and expressive. The added uncertainty introduced by the system noise error affects the KF results.

Deformation results for a selected group of points could be compared to expected deformation behavior which agreed in some cases but also gave information about unexpected deformation situations.

The vital importance of the link of the datum definition in the LHC tunnel to network points in the cavern has been identified in the dis-

cussion of the simulation example of the ATLAS cavern network. This link is foreseen to be realized by high precision measurement systems as described in Chapter 3. These systems have not been available during most of or possibly all of the detector installation period. The lack of these systems puts high demands on measurements passing through the survey galleries which are themselves poorly controlled due to bad geometry. Regular remeasurements of this network sections are recommended for verification until the datum definition link systems are available.

7

Conclusions

The surveying work in a project like the Large Hadron Collider (LHC) and related experiments at CERN is a crucial contribution to the project to be successful and the range of demands on surveying are diverse and challenging. Positioning and alignment of large physics detectors with respect to a particle beam line are very special applications in the field of engineering surveying in terms of accuracy requirements and working conditions.

In this thesis the problems of experimental cavern networks aligning and positioning various objects in the cavern with respect to the reference defined by the accelerator geometry have been discussed. Taking the ATLAS detector cavern as example, it could be seen that the two major problems of the network are cavern deformations and degrading network configurations. Deformations of the cavern walls and floors and of access structures inside the cavern are propagated onto the network geometry as the network points are directly linked to these.

The degradation of the network configuration is caused by the progress of the detector installation, increasingly limiting sightings between network points and disabling several network points.

It has been argued that classical deformation analysis methods based on congruency comparison between two epochs are not suited to adequately handle these problems, especially as the network configurations are very heterogenous in different epochs.

In Chapter 5 a kinematic model setup to represent the deforming network has been developed in the formalism of the Kalman Filter (KF). The kinematic interpretation of a geodetic network is a modern approach very well suited for high precision networks where the basic assumption of network point stability over long time periods is not appropriate. This interpretation has been newly introduced in the context of metrological networks for the alignment of accelerator and experimental detectors.

By handling a geodetic network in kinematic interpretation in a KF, it is very easy to accommodate changing network configurations, which are incomplete or even rank deficient in terms of insufficient datum definition. Improving the system description with additional data enables to profit in later epochs from better controlled and more reliable information of previous epochs. By implementing this idea, new in the context of the KF, the reliability in the network and measurements in this network can be ensured even if sections of the network cannot be accessed or points become unavailable.

An important property of the developed KF is its adaptive character. Adaptive filter methods usually can exploit large quantities of data sets to identify insufficiently or wrongly defined stochastic information in the filter process. In the case of surveying data in a geodetic network the number of data epochs is very small. The method of stochastic stabilization presented in Chapter 5.5 uses the KF system description derived from pervious data epochs and measurement data from a current epoch to iteratively identify correct stochastic information for the system noise as well as measurements. Thus changes in the system description (i.e. unexpected deformations) and gross errors in the observation data are identified by the adaptive KF in a parallel process which is iterated in each epoch of new measurement data. The combined application of this method to the processing of classical surveying data in a KF is newly introduced in this thesis. The implementation of the adaptive KF ensures that the common problem of filter divergence is avoided and that the kinematic system model representing a deforming geodetic network is correctly identified.

Three different data sets have been analyzed with an implementation of the presented algorithm:

A very simple simulation example has been used to show the general properties of the adaptive KF to handle a network in changing

configurations. It has been demonstrated that position and deformation estimates can be derived more efficiently and more accurately from the survey data compared to results of single epoch least-squares adjustment (SE LSQ) results and to the true simulated point coordinates. It could be seen that the KF estimates are less affected by bad network configurations, poorly controlled or uncontrolled observations and missing observations. KF estimates are also available for points which have not been measured in a certain epoch by means of the predicted values based on the system description. The corresponding, predicted error estimates give useful information about point accuracies for the planning of survey measurements. The effect of stochastic stabilization on the Kalman gain matrix has been illustrated using the example of deliberately introduced gross errors to the measurement data set and unexpected large point movements. It has been shown that the results of the adaptive KF can be improved in terms of accuracy and precision if additional information about point movements is included.

A simulation of the ATLAS design network has been used to analyze the network by means of reliability measures. A focus on the link of the datum definition to the remaining network identified these sections as critical to the whole network layout. The reliability in terms of measurement redundancy and precision of measurements providing this link is vital to accurate positioning. It has been shown that the reliability in the network observations is increased by the application of the KF compared to the SE LSQ method especially if the network configuration degrades or is rank deficient due to missing datum information.

The application of the adaptive KF method to real ATLAS cavern network data gathered over a period of three years has shown that the method can handle also very heterogenous and difficult data. Although the measurement information for the full 3D network is very sparse and the point movements not systematic the KF results compare favorably to the SE LSQ results for the majority of cases.

The implementation presented in this thesis will be continued to be used in the future for the analysis of network data from the ATLAS cavern as well as other LHC experiment zones at CERN. This includes the analysis of data gathered in the period until the start–up of LHC but also the shut–down periods during LHC operation. As these periods are very limited in terms of time, manpower and access possibilities, the planning of the network measurements to be carried out will be a

crucial task. The adaptive KF implementation can provide very useful information for this. The focus during shut–down periods will be on verifying and reproducing detector part positions within the cavern network. Time and resources for dedicated network measurements will be very limited. The application of the adaptive KF to the available data will give reliable position estimates for the survey tasks but also improve the system description of point deformations. It is expected that once installation activities cease in the cavern, network point deformations, especially of the access structures, will reduce and adopt a more systematic behavior than observed so far.

The method presented in this thesis is intended to be implemented as part of a geodetic least-squares adjustment software, like the software LGC developed and used by the CERN survey group, providing necessary corrections for the shape of the earth and to include more observation types than the implementation described here.

Implemented in such a way it can be used as a very powerful tool for many applications ranging from structural monitoring in engineering surveying (e.g. landslides, monitoring of dams or bridges, etc.), to other experimental metrology applications or high precision accelerators like the LHC or future machines. The kinematic interpretation of a network improves the knowledge of any network, which has to be maintained over a longer period of time and where deformations reaching the measurement precision over this time cannot be neglected. The adaptive KF as I have presented it in this thesis is the perfect tool to adjust measurements in such a network and to gain efficiently from all available information.

Appendix A

S-Transformation

In this section the method of S-transformation will be explained in more detail. The main reference is [44].

In a general datum definition i condition equations are part of the adjustment problem:

$$\mathbf{B}_i^T \hat{\mathbf{x}}_i = \mathbf{0} \tag{A.1}$$

thus the functional adjustment model is extended to

$$\begin{aligned} \mathbf{A}\hat{\mathbf{x}}_i &= \mathbf{l} + \mathbf{v} \\ \mathbf{B}_i^T \hat{\mathbf{x}}_i &= \mathbf{0}. \end{aligned} \tag{A.2}$$

In order to eliminate the rank deficiency of d in the adjustment model with u unknowns, $\mathbf{B}_i$ has to fulfill the following condition

$$\operatorname{rank}\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{B}_i^T \end{bmatrix}\right) = u \tag{A.3}$$

For the extended model (A.2) the system of normal equations becomes

$$\begin{bmatrix} \mathbf{N} & \mathbf{B}_i \\ \mathbf{B}_i^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_i \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{n} \\ \mathbf{0} \end{bmatrix}. \tag{A.4}$$

The solution is obtained as

$$\begin{aligned} \hat{\mathbf{x}}_i &= \mathbf{Q}_{11,i}\,\mathbf{n} && \text{with} \\ \mathbf{Q}_{xx,i} &= \mathbf{Q}_{11,i} && = \mathbf{Q}_i. \end{aligned} \tag{A.5}$$

$\hat{\mathbf{x}}_i$ and $\mathbf{Q}_i$ represent a special solution of $\mathbf{N}$ and depend on the conditions defined in $\mathbf{B}_i$. $\mathbf{Q}_{11,i}$ is obtained by inversion of the extended system of normal equations (A.4). It is shown in [44, p. 244] that

$$\mathbf{Q}_{11,i} = \left(\mathbf{N} + \mathbf{B}_i\mathbf{B}_i^T\right)^{-1} - \mathbf{G}\left(\mathbf{B}_i^T\mathbf{G}\right)^{-1}\left(\mathbf{G}^T\mathbf{B}_i\right)^{-1}\mathbf{G}^T, \tag{A.6}$$

where $\mathbf{G}$ is the matrix of orthonormal eigenvectors corresponding to the d eigenvalues $\lambda_i = 0$ of the original adjustment model $\mathbf{N}$.

The relationship of a datum defined by the condition matrix $\mathbf{B}_i$ to another datum based on conditions $\mathbf{B}_k$ can be obtained by using the properties of $\mathbf{Q}_{11,i} = \mathbf{Q}_i$.

$$\mathbf{N}\mathbf{Q}_i = \mathbf{I} - \mathbf{B}_i\left(\mathbf{G}^T\mathbf{B}_i\right)^{-1}\mathbf{G}^T. \tag{A.7}$$

By transposing this expression the matrix $\mathbf{S}_i$ can be defined

$$\mathbf{S}_i = \mathbf{Q}_i\mathbf{N} = \mathbf{I} - \mathbf{G}\left(\mathbf{B}_i^T\mathbf{G}\right)^{-1}\mathbf{B}_i^T. \tag{A.8}$$

Multiplication of a cofactor matrix $\mathbf{Q}_k$ that has been derived based on a datum definition $\mathbf{B}_k$ by $\mathbf{S}_i$ from the left and $\mathbf{S}_i^T$ from the right gives

$$\mathbf{S}_i\mathbf{Q}_k\mathbf{S}_i^T = \mathbf{Q}_i\mathbf{N}\mathbf{Q}_k\mathbf{N}\mathbf{Q}_i = \mathbf{Q}_i. \tag{A.9}$$

As $\mathbf{Q}_k$ is a generalized inverse of $\mathbf{N}$ and $\mathbf{G}^T$ is orthogonal to $\mathbf{N}$ it follows that $\mathbf{N}\mathbf{Q}_k\mathbf{N} = \mathbf{N}$ and $\mathbf{Q}_i\mathbf{N}\mathbf{Q}_i = \mathbf{Q}_i$. Thus

$$\mathbf{S}_i\mathbf{Q}_k\mathbf{S}_i^{(T)} = \mathbf{Q}_i. \tag{A.10}$$

which accomplishes the transfer from a datum defined by $\mathbf{B}_k$ to a datum defined by $\mathbf{B}_i$ as long as it is assured that the type of datum definition is the same.

Similarly for the solution vector $\hat{\mathbf{x}}$ we can derive

$$\mathbf{N}\hat{\mathbf{x}}_k = \mathbf{n}, \tag{A.11}$$

multiplication by $\mathbf{Q}_i$ gives

$$\begin{aligned} \mathbf{Q}_i\mathbf{N}\hat{\mathbf{x}}_k &= \mathbf{Q}_i\mathbf{n} \\ \mathbf{S}_i\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_i. \end{aligned} \tag{A.12}$$

Thus the S-transformation can be summarized as transforming a solution based on datum k to a datum definition i by the following equations:

$$\begin{aligned} \hat{\mathbf{x}}_i &= \mathbf{S}_i\hat{\mathbf{x}}_k \\ \mathbf{Q}_i &= \mathbf{S}_i\mathbf{Q}_k\mathbf{S}_i^T. \end{aligned} \tag{A.13}$$

Appendix B

Structure of the gain matrix K

Gain matrix $\mathbf{K}$, [3u × n] as defined in (5.19).

$$\mathbf{K}_i = \mathbf{P}_{i|i-1}\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i)^{-1} = \mathbf{P}_{i|i-1}\mathbf{H}_i^T\mathbf{D}_i^{-1}. \quad (5.117)$$

Splitting up the individual elements $\mathbf{P}_{i|i-1}$, $\mathbf{H}_i$ and $\mathbf{R}_i$ into partitions for position, velocity and acceleration states:

Covariance matrix of predicted state $\mathbf{P}_{i|i-1}$, [3u × 3u]:

$$\mathbf{P}_{i|i-1} = \begin{bmatrix} \mathbf{P}_{i|i-1\ pos|pos} & \mathbf{P}_{i|i-1\ pos|vel} & \mathbf{P}_{i|i-1\ pos|acc} \\ \mathbf{P}_{i|i-1\ vel|pos} & \mathbf{P}_{i|i-1\ vel|vel} & \mathbf{P}_{i|i-1\ vel|acc} \\ \mathbf{P}_{i|i-1\ acc|pos} & \mathbf{P}_{i|i-1\ acc|vel} & \mathbf{P}_{i|i-1\ acc|acc} \end{bmatrix}. \quad (B.1)$$

Measurement matrix $\mathbf{H}_i$, [n × 3u]:

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{H}_{i\ pos} & \mathbf{H}_{i\ vel} & \mathbf{H}_{i\ acc} \end{bmatrix}. \quad (B.2)$$

Covariance matrix of measurements $\mathbf{R}_i$, [n × n]:

$$\mathbf{R}_i = \mathrm{diag}(r_{i,jj}) = \begin{bmatrix} \ddots & & \\ & r_{i,jj} & \\ & & \ddots \end{bmatrix}. \tag{B.3}$$

Similarly the combined elements $\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T$ and $\mathbf{P}_{i|i-1}\mathbf{H}_i^T$ are split up:

$\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T$, [n × n]:

$$\begin{aligned} \mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T \;=& \mathbf{H}_{i\ pos}\mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T \;+\mathbf{H}_{i\ vel}\mathbf{P}_{i|i-1\ vel|vel}\mathbf{H}_{i\ pos}^T \\ &+\mathbf{H}_{i\ acc}\mathbf{P}_{i|i-1\ acc|pos}\mathbf{H}_{i\ pos}^T \;+\mathbf{H}_{i\ pos}\mathbf{P}_{i|i-1\ pos|vel}\mathbf{H}_{i\ vel}^T \\ &+\mathbf{H}_{i\ vel}\mathbf{P}_{i|i-1\ vel|vel}\mathbf{H}_{i\ vel}^T \;+\mathbf{H}_{i\ acc}\mathbf{P}_{i|i-1\ acc|vel}\mathbf{H}_{i\ vel}^T \\ &+\mathbf{H}_{i\ pos}\mathbf{P}_{i|i-1\ pos|acc}\mathbf{H}_{i\ acc}^T \;+\mathbf{H}_{i\ vel}\mathbf{P}_{i|i-1\ vel|acc}\mathbf{H}_{i\ acc}^T \\ &+\mathbf{H}_{i\ acc}\mathbf{P}_{i|i-1\ acc|acc}\mathbf{H}_{i\ acc}^T. \end{aligned} \tag{B.4}$$

$\mathbf{P}_{i|i-1}\mathbf{H}_i^T$, [3u × n]:

$$\mathbf{P}_{i|i-1}\mathbf{H}_i^T = \left[\begin{array}{c} \mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T + \\ \mathbf{P}_{i|i-1\ pos|vel}\mathbf{H}_{i\ vel}^T + \\ \mathbf{P}_{i|i-1\ pos|acc}\mathbf{H}_{i\ acc}^T \\ \hline \mathbf{P}_{i|i-1\ vel|pos}\mathbf{H}_{i\ pos}^T + \\ \mathbf{P}_{i|i-1\ vel|vel}\mathbf{H}_{i\ vel}^T + \\ \mathbf{P}_{i|i-1\ vel|acc}\mathbf{H}_{i\ acc}^T \\ \hline \mathbf{P}_{i|i-1\ acc|pos}\mathbf{H}_{i\ pos}^T + \\ \mathbf{P}_{i|i-1\ acc|vel}\mathbf{H}_{i\ vel}^T + \\ \mathbf{P}_{i|i-1\ acc|acc}\mathbf{H}_{i\ acc}^T \end{array}\right] \tag{B.5}$$

The predicted state covariance matrix $\mathbf{P}_{i|i-1}$, [3u × 3u] is an important element. It is defined as

$$\mathbf{P}_{i|i-1} = \mathbf{\Phi}(i,i-1)\mathbf{P}_{i-1|i-1}\mathbf{\Phi}^T(i,i-1)+\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1), \quad \text{(B.6)}$$

and is composed of $\mathbf{P}_{i-1|i-1}$, $\mathbf{\Phi}(i,i-1)$, $\mathbf{\Gamma}(i-1)$ and $\mathbf{Q}_{i-1}$:

$\mathbf{P}_{i-1|i-1}$, [3u × 3u] which is the covariance matrix of the preceding state update and is composed as follows:

$$\mathbf{P}_{i-1|i-1} = \left[\begin{array}{c|c|c} \mathbf{P}_{i-1|i-1\ pos|pos} & \mathbf{P}_{i-1|i-1\ pos|vel} & \mathbf{P}_{i-1|i-1\ pos|acc} \\ \hline \mathbf{P}_{i-1|i-1\ vel|pos} & \mathbf{P}_{i-1|i-1\ vel|vel} & \mathbf{P}_{i-1|i-1\ vel|acc} \\ \hline \mathbf{P}_{i-1|i-1\ acc|pos} & \mathbf{P}_{i-1|i-1\ acc|vel} & \mathbf{P}_{i-1|i-1\ acc|acc} \end{array}\right], \quad \text{(B.7)}$$

the transition matrix $\mathbf{\Phi}$, [3u × 3u] as defined in (5.51) or (5.52):

$$\mathbf{\Phi}(i,i-1) = \begin{bmatrix} \mathbf{I} & \Delta t \cdot \mathbf{I} & \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \Delta t \cdot \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \text{(B.8)}$$

as well as $\mathbf{\Gamma}(i-1)$, [3u × u] as defined in (5.55)

$$\mathbf{\Gamma}(i-1) = \begin{bmatrix} \frac{1}{6}\Delta t^3 \cdot \mathbf{I} \\ \frac{1}{2}\Delta t^2 \cdot \mathbf{I} \\ \Delta t \cdot \mathbf{I} \end{bmatrix}, \quad \text{(B.9)}$$

and the covariance matrix $\mathbf{Q}_{i-1}$, [u × u] of the system noise $\mathbf{w_{i-1}}$, as defined before in (5.65):

$$\mathbf{Q}_{i-1} = \mathrm{diag}(\mathbf{q_{i-1}}). \quad \text{(B.10)}$$

The product $\mathbf{\Phi}(i,i-1)\mathbf{P}_{i-1|i-1}\mathbf{\Phi}(i,i-1)^T$, [3u × 3u] thus follows as:

$$\mathbf{\Phi}(i,i-1)\mathbf{P}_{i-1|i-1}\mathbf{\Phi}(i,i-1)^T =$$

$$\left[\begin{array}{c|c|c}
\begin{array}{l}\mathbf{P}_{i-1|i-1\ pos|pos}+\\ 2\Delta t\mathbf{P}_{i-1|i-1\ pos|vel}+\\ \Delta t^2\mathbf{P}_{i-1|i-1\ pos|acc}+\\ \Delta t^2\mathbf{P}_{i-1|i-1\ vel|vel}+\\ \Delta t^3\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \frac{1}{4}\Delta t^4\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} &
\begin{array}{l}\mathbf{P}_{i-1|i-1\ pos|vel}+\\ \Delta t\mathbf{P}_{i-1|i-1\ pos|acc}+\\ \Delta t\mathbf{P}_{i-1|i-1\ vel|vel}+\\ \frac{3}{2}\Delta t^2\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \frac{1}{2}\Delta t^3\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} &
\begin{array}{l}\mathbf{P}_{i-1|i-1\ pos|acc}+\\ \Delta t\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \frac{1}{2}\Delta t^2\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} \\
\hline
\begin{array}{l}\mathbf{P}_{i-1|i-1\ pos|vel}+\\ \Delta t\mathbf{P}_{i-1|i-1\ pos|acc}+\\ \Delta t\mathbf{P}_{i-1|i-1\ vel|vel}+\\ \frac{3}{2}\Delta t^2\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \frac{1}{2}\Delta t^3\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} &
\begin{array}{l}\mathbf{P}_{i-1|i-1\ vel|vel}+\\ 2\Delta t\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \Delta t^2\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} &
\begin{array}{l}\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \Delta t\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} \\
\hline
\begin{array}{l}\mathbf{P}_{i-1|i-1\ pos|acc}+\\ \Delta t\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \frac{1}{2}\Delta t^2\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} &
\begin{array}{l}\mathbf{P}_{i-1|i-1\ vel|acc}+\\ \Delta t\mathbf{P}_{i-1|i-1\ acc|acc}\end{array} &
\mathbf{P}_{i-1|i-1\ acc|acc}
\end{array}\right] \tag{B.11}$$

and the product $\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1)$, [3u × 3u] as:

$$\mathbf{\Gamma}(i-1)\mathbf{Q}_{i-1}\mathbf{\Gamma}^T(i-1) =$$

$$\left[\begin{array}{c|c|c}
\frac{1}{36}\Delta t^6\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} & \frac{1}{12}\Delta t^5\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} & \frac{1}{6}\Delta t^4\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} \\
\hline
\frac{1}{12}\Delta t^5\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} & \frac{1}{4}\Delta t^4\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} & \frac{1}{2}\Delta t^3\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} \\
\hline
\frac{1}{6}\Delta t^4\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} & \frac{1}{2}\Delta t^3\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I} & \Delta t^2\cdot\mathrm{diag}(\mathbf{q_{i-1}})\cdot\mathbf{I}
\end{array}\right] \tag{B.12}$$

In the special case of $\mathbf{H}_{i\ vel} = \mathbf{H}_{i\ acc} = \mathbf{0}$, for all i simplifications arise:

$$\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T = \mathbf{H}_{i\ pos}\mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T \tag{B.13}$$

$$\mathbf{P}_{i|i-1}\mathbf{H}_i^T = \begin{bmatrix} \mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T \\ \hline \mathbf{P}_{i|i-1\ vel|pos}\mathbf{H}_{i\ pos}^T \\ \hline \mathbf{P}_{i|i-1\ acc|pos}\mathbf{H}_{i\ pos}^T \end{bmatrix} \tag{B.14}$$

Summarizing, the result for the gain matrix $\mathbf{K}_i$ is:

$$\begin{aligned} \mathbf{K}_i &= \mathbf{P}_{i|i-1}\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_{i|i-1}\mathbf{H}_i^T + \mathbf{R}_i)^{-1} = \\ &= \begin{bmatrix} \mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T \\ \hline \mathbf{P}_{i|i-1\ vel|pos}\mathbf{H}_{i\ pos}^T \\ \hline \mathbf{P}_{i|i-1\ acc|pos}\mathbf{H}_{i\ pos}^T \end{bmatrix} \cdot (\mathbf{H}_{i\ pos}\mathbf{P}_{i|i-1\ pos|pos}\mathbf{H}_{i\ pos}^T + \mathbf{R}_i)^{-1} \end{aligned} \tag{B.15}$$

List of References

[1] B. D. O. Anderson and J. B. Moore *Optimal Filtering.* Prentice Hall, Inc., Englewood Cliffs, N.J., 1979.

[2] ATLAS Collaboration. *ATLAS Technical Design Report (TDR) 13.* CERN, Geneva, Switzerland, 1999.

[3] ATLAS Collaboration. *The ATLAS experiment.* Project webpage supported by CERN, `http://atlas.ch`.

[4] W. Baarda. *A Testing Procedure for Use in Geodetic Networks.* Netherlands geodetic Comission, Vol. 2/5, Delft, 1968.

[5] F. Becker, W. Coosemans, M. Jones. *Consequences of Perturbations of the Gravity Field on HLS Measurements.* In: Proceedings of the 7th International Workshop on Accelerator Alignment, SPring-8, Japan, November 2002.

[6] R. G. Brown and P.Y.C. Hwang. *Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions.* John Wiley & Sons, 3rd edition, New York, 1997.

[7] F. Butin. *Support Structures System Analysis: Results of the Second Iteration.* Internal report CERN , ATLAS Project ATC-T-ER-0002, EDMS N° 100132, May 2000.

[8] M.A.R. Cooper *Control surveys in civil engineering.* Collins Professional and Technical Books, London, 1987.

[9] W. Coosemans and H. Mainaud. *Pre-Alignement of CLIC using the double-wire method.* Contribution to the Particle Accelerator Conference, Vancouver, Canada, May 1997.

[10] W. Coosemans, H. Mainaud Durand, A. Marin, J-P Quesnel. *The Alignment of the LHC Low Beta Triplets: Review of Instrumentation and Methods.* In: Proceedings of the 7th International Workshop on Accelerator Alignment, SPring-8, Japan, November 2002.

[11] C. Cheval. *Verification of the ATLAS Surrounding Structure (HS).* Technical report CERN-TIS&TE TIS/TE/MC/3-8, EDMS N° 327212, October 2001.

[12] Electricité de France (EDF). *Excavation du stross de la caverne UX15 - Analyse tridimensionnelle du comportement des ouvrages souterrains durant le creusement du stross de la caverne UX15.* Technical note by contractor EDF, IH-CERN-VAR62-00004-A-BPE, 2002.

[13] M. Forsblom. *Softwares for deformation analysis in the ATLAS cavern network.* Internal report CERN, September 2000.

[14] J.C. Gayde, A. Herty, H. Mainaud Durand, C. Lasseur. *Combined Levelling Systems for the Vertical Monitoring of a Large Physics Experiment.* In: H. Kahmen and A. Chrzanowski (Eds.) Proc. 3^{rd} IAG Symposium on Geodesy for Geotechnical and Strcutural Engineering and 12^{th} International Symposium on Deformation Measurements (CD), Baden, Austria, 2006.

[15] A. Gelb (Ed.). *Applied Optimal Estimation.* The M.I.T. Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, and London, England, 1974.

[16] E. W. Grafarend and B. Schaffrin. *Ausgleichsrechnung in linearen Modellen.* BI-Wissenschaftsverlag, Mannheim, 1993.

[17] C. Guitton. *Etudes des ouvrages souterrains ATLAS.* In: 1st ST Workshop, Chamonix, France, 3-6 February 1998 - pages 89-100.

[18] R. Hawkings. *Survey of the ATLAS ID.* ATLAS Project Note, ATLAS Project ATL-IC-IN-0001, EDMS N° 114641, June 2000.

[19] O. Heunecke. *Zur Identifikation und Verifikation von Deformationsprozessen mittels adaptiver KALMAN-Filterung (Hannoversches Filter).* Nr. 208 in Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Hannover, 1995.

[20] A.-K. Hocq and D. Blunier. *Finite Element Calculation of the HO Structure.* Calculation note CERN, ATLAS Project ATL-HO-EN-0004, EDMS N° 113780, August 2001.

[21] A.H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, New York, 1970.

[22] M. Jones. *Geodetic and Astronomical Reference & Coordinate Systems.* CERN Internal Note, EDMS N° 107906, November 1999.

[23] M. Jones. *Geodetic Definition (Datum Parameters) of the CERN COORDINATE SYSTEM.* CERN Internal Note, EDMS N° 107981, February 2000.

[24] M. Jones, M. Mayoud and A.Wiart. *Geodetic Parameterisation of the CNGS Project.* In: Proceedings of the 7th International Workshop on Accelerator Alignment, SPring-8, Japan, 2002.

[25] T. Kailath. *An Innovations Approach to Least-Squares Estimation. Part I: Linear Filtering in Additive White Noise.* IEEE Transactions on AUTOMATIC CONTROL, VOL. AC-13, NO. 6, December 1968.

[26] R.E. Kalman. *A New Approach to Linear Filtering and Prediction Problems.* Reprinted in: H.W. Sorenson (Ed.): Kalman Filtering: Theory and Application, Transactions of the ASME, Journal of Basic Engineering, Series 82D, pp. 35-45, March 1960.

[27] K. R. Koch. *Parameterschätzung und Hypothesentests.* Ferd. Dümmlers Verlag, 3. Auflage, Bonn, 1997.

[28] T. Krarup, J. Juhl and K. Kubik. *Götterdämmerung over least squares adjustment.* In: 14^{th} Congress of the International Society of Photogrammetry, Vol B3, Hamburg, 1980: pages 369 - 371.

[29] C. Lasseur. *Alignment of the LHC machine and experiments - Part 2) The experiments.* Engineering/Technical Note CERN, Lemic 27 November 2001, EDSM N° 330354 (https://edms.cern.ch/document/330354/1), November 2001.

[30] C. Lasseur, M.Mayoud, D.Missiaen and J.P. Quesnel. *Géométrie du LHC: Points caractéristiques, Formules de Transformation.* LHC Project Note 95, EDSM N° 108124 (https://edms.cern.ch/document/108124), June 1997.

[31] LGC. *Logiciel Général de Compensation, Version 3.2.1.* Software manual, CERN, Switzerland, 2002.

[32] A. Lippitsch. *An Alternative Approach in Metrological Network Deformation Analysis Employing Kinematic and Adaptive Methods.* In: H. Kahmen and A. Chrzanowski (Eds.) Proc. 3^{rd} IAG Symposium on Geodesy for Geotechnical and Strcutural Engineering and 12^{th} International Symposium on Deformation Measurements (CD), Baden, Austria, 2006.

[33] F. Löffler u.a. *Maschinen- und Anlagebau.* in M. Möser, G. Müller, H. Schlemmer und H. Werner (Hrsg.): Handbuch Ingenieurgeodäsie, 2., völlig neu bearb. und erw. Auflage, Herbert Wichmann Verlag, Heidelberg, 2002.

[34] H. Mainaud Durand, A. Herty, A. Marin, A. Mathieu, T. Rénaglia. *Status of the Alignment of the LHC Low beta Quadrupoles.* In: Proceedings of the 8th International Workshop on Accelerator Alignment, CERN, Geneva, Switzerland, October 2004.

[35] P.S. Maybeck *Stochastic Models, Estimation, and Control - Volume 1.* Acadedemic Press, New York, 1979.

[36] P.S. Maybeck *Stochastic Models, Estimation, and Control - Volume 2.* Acadedemic Press, New York, 1982.

[37] T.P. McGarty *Stochastic systems and state estimation.* John Wiley & Sons, New York, 1974.

[38] R.K. Mehra. *On the Identification of Variances and Adaptive Kalman Filtering.* IEEE Transactions on AUTOMATIC CONTROL, VOL. AC-15, NO.2, pp. 175-184, April 1970.

[39] R.K. Mehra. *Approaches to Adaptive Filtering.* IEEE Transactions on AUTOMATIC CONTROL, VOL. AC-17, NO.2, pp. 693-698, October 1972.

[40] P. Meissl. *Zusammenfassung und Ausbau der inneren Fehlertheorie eines Punkthaufens.* Reihe A - Heft Nr. 61, Deutsche Geodätische Kommision bei der Bayerischen Akademie der Wissenschaften, München,1969.

[41] G. Minkler and J. Minkler. *Theroy and Application of Kalman Filtering.* Magellan Book Company, Palm Bay, 1993.

[42] T. Möller and B. Trumbore. *Fast, minimum storage ray-triangle intersection.* Journal of graphics tools, 2(1): 21-28, 1997.

[43] W. Niemeier. *Zur Kongruenz mehrfach beobachteter geodätischer Netze.* Nr. 88 in Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Hannover, 1979.

[44] W. Niemeier. *Ausgleichsrechnung: eine Einführung für Studierende und Praktiker des Vermessungs- und Geoinformationswesens.* de Gryter, Berlin, New York, 2002.

[45] H.B. Papo and A. Perelmutter. *Pre-zero-epoch covariance matrix in sequential analysis of deformations.* Bulletin Géodésique 58 (1984): pp. 75-83, 1984.

[46] R.J.H. Parkin. *Design and construction of the tension ties for the UX15 cavern vault.* Presented at the 5th ST Workshop, Echenevex, France, 28-30 January 2002, ST-Note-2002-018.

[47] H. Pelzer. *Zur Analyse geodätischer Deformationsmessungen.* Reihe C: Dissertationen - Heft Nr. 164, Deutsche Geodätische Kommision bei der Bayerischen Akademie der Wissenschaften, München,1971.

[48] H. Pelzer (Hrsg.). *Geodätische Netze in der Landes- und Ingenieurvermessung II.* Vorträge des Kontaktstudiums 1985, Verlag Konrad Wittwer KG, Stuttgart,1985.

[49] H. Pelzer. *Deformationsuntersuchungen auf der Basis kinematischer Bewegungsmodelle.* Allgemeine Vermessungs-Nachrichten,

Herbert Wichmann Verlag GmbH Stuttgart, 94. Jg. Heft 2, S. 49-62, Februar 1987.

[50] H. Rammer. *Two new caverns for LHC experiments : ATLAS and CMS* In: 1st ST Workshop, Chamonix, France, 3-6 February 1998 - pages 82-88.

[51] H. Rammer. *Novel technique for the UX15 cavern vault support system.* In: 3rd ST Workshop, Chamonix, France, 25-28 January 2000 - pages 301-305.

[52] H. Rammer. *LEP tunnel movements at Point 1 caused by LHC civil engineering.* In: 4th ST Workshop, Chamonix, France, 30 January - 2 February 2001 - pages 229-232.

[53] H. Rammer. Personal communication March 2003.

[54] RoboBAT. *Manuel d'Utilisation ROBOT 97 Version 12.5.* Software manual, © RoboBAT, Meylan, France, 1997.

[55] L. Sachs. *Angewandte Statistik - Anmwendung statistischer Methoden.* Springer-Verlag, 9. Auflage, Berlin, Heidelberg, 1999.

[56] M.A. Salzmann. *Some Aspects of Kalman Filtering.* Technical Report 140, Department of Surveying Engineering, University of New Brunswick, Fredericton, Canada, 1998.

[57] K.-W. Schrick (Hrsg.). *Anwendungen der Kalman-Filter-Technik: Anleitung und Beispiele.* Methoden der Regelungstechnik, R. Oldenbourg Verlag, München, Wien, 1977.

[58] H.W. Sorenson (Ed.). *Kalman Filtering: Theory and Application.* IEEE Press, The Institute of Electrical and Electronics Engineers, Inc., New York, 1985.

[59] G. Strang and K. Borre. *Linear Algebra, geodesy, and GPS.* Wellesley-Cambridge Press, Wellesley, 1997.

[60] W. M. Welsch and O. Heunecke. *Models and Terminology for the Analysis of Geodetic Monitoring Observations.* - Official Report of the Ad-Hoc Committee of FIG Working Group 6.1, In: Proceedings of the 10th FIG International Symposium on Deformation

Measurements, 19 - 22 March 2001 Orange, California, USA, pages 390-412, 2001.

[61] W. Welsch, O. Heunecke and H. Kuhlmann. *Auswertung geodästischer Überwachungsmessungen.* in M. Möser, G. Müller, H. Schlemmer und H. Werner (Hrsg.): Handbuch Ingenieurgeodäsie, Herbert Wichmann Verlag, Heidelberg, 2000.

[62] A. Wieser. *Robust and fuzzy techniques for parameter estimation and quality assessment in GPS.* Dissertation Technische Universität Graz 2001, erschienen im Shaker Verlag, Aachen 2002.

[63] A. Wieser, M.G. Petovello and G. Lachapelle. *Failure Scenarios to be Considered with Kinematic High Precision Relative GNSS Positioning.* ION GNSS 17th International Technical Meeting of the Satellite Division, Long Beach, CA, 2004: 1448 - 1459.

Curriculum vitae

18.7.1976	Born in Graz/Austria as daughter of Gerhilt and Ing. Erich Lippitsch
17.6.1994	Final Exam (Matura) at WIKU BRG Graz, Austria
1994–2000	Study of Surveying at Graz University of Technology, Austria
1997–1998	ERASMUS free mover student at Nottingham Trent University, UK
24.6.1998	Graduation BEng at Nottingham Trent University Final Year Project thesis on 'Real Time On-Site Prediction of GPS-Observations'
22.3.2000	Graduation (Dipl.-Ing.) in Surveying at TU Graz Specialization in Engineering Surveying Diploma thesis on 'Quality Control for GPS RTK Measurements in Engineering Geodesy'
2000–2001	Research Assistant at the Institute of Engineering Geodesy and Measurement Systems, Prof. F.K. Brunner, TU Graz
since 2001	Doctoral Student in TS/SU/EM and Fellow in PH/DT2 at CERN, CH. Working in the Experiment Metrology section of TS/SU, Christian Lasseur