NTNU
Norwegian University of
Science and Technology

A heavy tailed statistical model applied in anti-collision calculations for petroleum wells

Tony Gjerde

Master of Science in Physics and Mathematics
Submission date: June 2008
Supervisor: Jo Eidsvik, MATH
Co-supervisor: Bjørn Torstein Bruun, StatoilHydro

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

Analyses of measurements from directional surveys in directional drilling indicate that the use of heavy tailed statistical distributions is needed. A heavy tailed statistical model is applied on the position uncertainty of a wellbore. Anti-collision calculations are performed with both the heavy-tailed model and the model which is currently used as a petroleum industry standard, and the differences between the models is analyzed.

Assignment given: 09. January 2008
Supervisor: Jo Eidsvik, MATH

NTNU
Norwegian University of Science and Technology

Faculty of Information Technology, Mathematics and Electrical Engineering

MASTER'S THESIS

FOR

STUD.TECHN. Tony Gjerde

FACULTY OF INFORMATION TECHNOLOGY,
MATHEMATICS AND ELECTRICAL ENGINEERING

NTNU

Date due: June 4, 2008

Discipline: Statistics

Title: " A heavy tailed statistical model applied in anti-collision calculations for petroleum wells"

Purpose of the work: Analyses of measurements in directional drilling indicate that the use of heavy tailed statistical distributions is needed. We apply a heavy tailed statistical model on the position uncertainty of a wellbore. We perform anti-collision calculations with both the heavy tailed model and the model which is currently used as a petroleum industry standard, and analyze the differences in the results of these models.

This diploma thesis is to be carried out at the Department of Mathematical Sciences under guidance of Associate Professor Jo Eidsvik.

Trondheim, January 9, 2008.

Trond Digernes Instituttleder Dept. of Mathematical Sciences	Jo Eidsvik Associate Professor Dept. of Mathematical Sciences

Preface

This Master's thesis is written the spring semester 2008 at Norwegian University of Science and Technology (NTNU), Faculty of Information Technology, Mathematics and Electrical Engineering (IME), Department of Mathematical Sciences. The Master's thesis is the last and final subject at the study program Master of Science (Sivilingeniør) in Industrial Mathematics, with specialization in statistics. The name of the subject is "TMA4905 - Statistikk, masteroppgave", and it is normed to 30 university credit points. The Master's thesis is written in collaboration with StatoilHydro, which provided me with the problem of this thesis. An agreement regarding the intellectual property rights, the rights to publish, the commercial and financial rights of the contents in this Master's thesis has been signed and distributed between the involved parties.

Associate Professor Jo Eidsvik has been my supervisor at NTNU during the last year. We have had an excellent dialogue, and he has always had supportive advices for me when I have needed it the most. My supervisors in StatoilHydro, Bjørn Torstein Bruun and Erik Nyrnes, have been professional advisors to me in subjects which I am no claimed expert in. I thank all three for the time and effort they have given to support me during my work on the Master's thesis.

Trondheim, June 2008

Tony Gjerde

Abstract

Anti-collision calculations are done during the planning of a new petroleum well. These calculations are required in order to control the risk of having a well-collision, which is an unwanted event at any cost. The risk of having a well-collision is closely related to the position uncertainty both of the well that is planned and of the existing wells in the given region. Earlier literature has indicated that the distribution of the position errors are more heavy-tailed than a normal distribution, which leads to the question whether the current methods are accurate enough. The currently used industry standard calculates the standard deviation of the centre to centre distance by an approximation, and assumes that the centre to centre distance is normally distributed. In this thesis we use a heavy-tailed Normal Inverse Gaussian (NIG) distribution for the declination error source in MWD magnetic directional surveying, which lead to a position uncertainty that is heavy-tailed relative to the multivariate normal distribution. The parameters of the NIG-distribution are estimated from processed magnetic field data from the Tromsø geomagnetic observation station. The NIG-distribution requires the use of Monte Carlo simulations in order to apply the currently used industry approach. Other error sources are also included in the error model to give a more realistic position uncertainty.

Three different anti-collision cases demonstrate the differences in using the NIG error model and the normal error model. We compare the simulation based results against the currently used methodology. The results are very dependent on the well geometries. The results differ significantly, and the NIG error model is the most conservative distribution in most cases, with respect to whether a wellplan should be realized or not. However, there are cases where a normally distributed declination error gives more conservative decisions than the NIG-distribution. As an alternative to change the distribution of the declination error, we propose two corrective actions to improve the existing anti-collision methodology. One action is to exchange one of the approximations in the current methodology with simulations or analytical computations. The other action is to correct for bias in the expected position, which is caused by the NIG error model.

Contents

1 Introduction **1**
1.1 Introduction 1
1.2 Overview 3

2 MWD magnetic directional surveying **5**
2.1 Introduction 5
2.2 Coordinate systems 7
2.3 Magnetic reference system 8
2.4 Gravity reference system 9
2.5 Connection between the instrument based- and the magnetic reference coordinate system 9
2.6 Minimum curvature methods 12

3 Statistical distributions **15**
3.1 Introduction 15
3.2 Statistical moments 15
3.3 The normal distribution 17
3.4 The normal inverse gaussian distribution 18
  3.4.1 Density function 18
  3.4.2 Parameter estimation by EM-algorithm 20
  3.4.3 Simulation 23
3.5 Other skewed and heavy-tailed distributions 23

4 Normally distributed position uncertainty in MWD magnetic directional surveying **25**
4.1 Introduction 25
4.2 Error model 26
  4.2.1 Accelerometer bias errors 26
  4.2.2 Accelerometer scale errors 28
  4.2.3 Magnetometer bias errors 29
  4.2.4 Magnetometer scale errors 30
  4.2.5 Depth errors 30
  4.2.6 Declination errors 31
  4.2.7 Sag error 32
  4.2.8 Axial magnetism error in drillstring 32
  4.2.9 Excluded errors 32
4.3 Error propagation theory 33

5 The normal inverse gaussian distribution in MWD magnetic directional surveying 39
5.1 Introduction 39
5.2 Specification of NIG-parameters for the declination error 40
5.2.1 Symmetric NIG-distributed declination error 42
5.2.2 Skewed NIG-distributed declination error 42
5.3 Approximative density function 43
5.4 Simulating the position uncertainty with the NIG error model . . 45
5.5 Some analytical calculations 47
5.6 Approximations for special well geometries 48
5.7 Comparing the discretized approximative density function with the normal error model 49

6 Anti-collision calculations 53
6.1 Introduction 53
6.2 Scan methods 54
6.2.1 3D Closest approach 55
6.2.2 Travelling cylinder 55
6.2.3 Modified travelling cylinder 55
6.2.4 Horizontal plane 56
6.3 Comparison of the scanning methods 56
6.4 Anti-collision hypothesis testing 57
6.4.1 Hypothesis test on the distance D: Test I 58
6.4.2 Hypothesis test on the empirical distance D: Test II . . . 59
6.5 The probability of collision 60

7 Results of the anti-collision calculations 63
7.1 Introduction 63
7.2 Presentation of anti-collision simulations 63
7.3 Anti-collision Case I 66
7.4 Anti-collision Case II 72
7.5 Anti-collision Case III 77
7.6 General results 81

8 Closing remarks 85
8.1 Conclusions 85
8.2 Further work 87

A Proof: Linearity formulas of the normal inverse gaussian distribution I

B Proof: The normal inverse gaussian distribution as a normal variance mean-mixture III

C Normal approximation in anti-collision hypothesis testing V

D An alternative hypothesis test in anti-collision calculations VII

List of Tables

4.1 Overview of the different error sources used in this thesis. All the error sources in this table are considered to be uncorrelated. The way these error sources affect MWD-surveys is shown in the propagation mode column. (R: Random, S: Systematic, G: Global). The standard deviations in the table are from [25]. . . . 27

5.1 A summary of the choice of NIG-parameters for the two NIG-distributions of the declination error. The symmetric NIG-distribution is partly obtained from [22], while the skewed NIG-distribution is obtained from real processed data [10]. 43

6.1 Summary of the advantages (+) and disadvantages (−) of the different scanning methods. 56

7.1 Overview of the different separation factors that are analyzed in the anti-collision calculations. 64
7.2 The empirical eigenvalues (top) and eigenvectors (below) for the reference well and offset well of anti-collision case I. 66
7.3 The empirical eigenvalues (top) and eigenvectors (below) of anti-collision case II. 72
7.4 The empirical eigenvalues (top) and eigenvectors (below) for the reference well and offset well of anti-collision case III. 77

List of Figures

2.1 Definition of azimuth angle A and the inclination angle I in the NEV coordinate-system. The measured depth D is the "along-hole distance". 6

2.2 Definition of toolface angle τ and the high-side direction in the instrument based coordinate-system, which is defined by the way the sensors are mounted. The z-direction is the down-hole direction, and it is orthogonal to the plane spanned by x and y. The high-side direction is defined as the negative V-direction (upward-direction) perpendicular to the z-axis. 7

2.3 Definition of the magnetic field vector in the NEV coordinate-system. The field strength defines the length of B, and the direction of B is defined by the magnetic dip angle θ and the magnetic declination angle δ. 8

2.4 Schematic illustration of the notation used in the minimum curvature method. p_{k-1}, $p^\star$ and p_k are positions along the wellpath. MWD-measurements are made in p_{k-1} and p_k, where t_{k-1} and t_k are the direction vectors. $p^\star$ is the position at a desired measured depth $D^\star$, which we must interpolate the position to. 11

3.1 Three different probability density functions. The dashed red line is a standard normal distribution, the green line is a symmetric distribution with heavy tails, and the blue line is a right-skewed distribution. 16

5.1 Choice of NIG-distribution for two of the error sources. The dashed red lines are the normal distributions proposed in the error model in Section 4.2. Left: $\varepsilon_1 \sim NIG(0, 0.03, 1000, 0.07)$. Right: $\varepsilon_7 \sim NIG(-196, 2800000, 1000, 0.07)$. 41

5.2 The two proposed NIG-distributions for the declination error. The blue line is the symmetric NIG-distribution, and the green line is the skewed NIG-distribution. The histogram is the centre of the processed dataset Tromsø 2001, as the range of the processed dataset is 0.26 radians. The dashed red line is the normal distribution. The NIG-parameters of the models are shown in Table 5.1. 43

5.3 An example of a simulation plot where the declination error is NIG-distributed and the other errors are normally distributed. . 46

5.4 Contour plots of the approximative marginal distributions of the NIG error model. The approximative density function is evaluated on a 3D-grid and the 2D-marginals are computed by summing over the last dimension. The contour lines are manually chosen. We see that the contours of the approximative density function is non-normal, especially in the tails. 51

5.5 The NE-marginal of the corresponding multivariate normal distribution from the normal error model of anti-collision case II. The figure is comparable with the upper plot in Figure 5.4, and the contour lines have the same values as in Figure 5.4. 52

6.1 Schematic illustration of the different scanning methods which are commonly used for classifying the closest point in an offset well. The filled points in the reference well are points where anti-collision calculations have been performed. p_1 is the candidate point of the 3D Closest approach method, p_2 is the candidate point of the travelling cylinder method, and p_3 is the candidate point of the horizontal plane method. 54

6.2 Schematic illustration of the calculated coordinate u_0 in the reference well, the calculated coordinate v_0 in the offset well, and the calculated minimum centre to centre distance D_0. 57

6.3 Histogram of $D \sim f_D(d)$ with $E(D) = \mu_D$ and $\text{Var}(D) = \sigma_D^2$ from a typical simulation (left). Histogram of $\frac{D-\mu_D}{\sigma_D}$ (right). The empirical 99.5%-quantile k_2 is found from the distribution of the standardized distance in hypothesis test II. 59

7.1 An example of an anti-collision calculation based on simulations. This simulation shows two synthetic wells that have a high risk of intersecting. All separation factors are much lower than 1, which indicates that this reference well would never be drilled after this plan. 65

7.2 Anti-collision case I. A synthetic northern directed well with constant inclination ($I = 45$) is close to intersecting with a real offset well. We see that $\omega_1 > 1$, signalizing that there is no significant risk of well-collision given that the normal error model is correct, while $\omega_3 < 1$, signalizing the opposite for the NIG error model. . 67

7.3 A thorough analysis of the anti-collision case I, where the straight reference well ($I = 45, A = 0$) is shifted on a grid in the NE-plane and the anti-collision calculations are performed against the point ($D = 3200, I = 7.5, A = 274$) in the offset well. Contour plot of ω_1 is shown in the upper figure, ω_3 is shown in the middle figure, and the relative difference $\frac{\omega_3-\omega_1}{\omega_3}$ is shown in the lower figure. We see that there are large differences between ω_1 and ω_3, where ω_3 is the most conservative separation factor with respect to the hypothesis tests. 69

7.4 Contour plots of the empirical quantile k_3 (upper), the empirical estimate $\hat{\sigma}_{D,1}$ (middle) and $\hat{\sigma}_{D,3}$ (lower) for the anti-collision case I where the reference well is shifted on a grid in the NE-plane. The dashed line indicates where $\omega_1 = 1$. 70

7.5 Contour plots of the normal simulated ω_2 (upper), and the relative difference $\frac{\omega_2-\omega_1}{\omega_2}$ (centre), and the relative difference $\frac{\omega_2-\omega_3}{\omega_2}$ (lower) for anti-collision case I. 71

7.6 Anti-collision case II. We perform the anti-collision analyses of two identical parallel wells at $D = 5000$ meters measured depth. The inclination angle is approximately 66 degrees, and the azimuth angle is approximately 305 degrees. Both wells have identical position uncertainty because of the equal well geometries, so only one of them are shown in this figure. 73

7.7 A thorough analysis of the anti-collision case II where one of the wells are shifted in a grid in the NE-plane. Contour plot of ω_1 (upper), ω_3 (middle), and the relative difference $\frac{\omega_3-\omega_1}{\omega_3}$ (lower). We see that the relative difference is both greater and smaller than 0, which gives no certain answer for which error model is the most conservative for this case. 74

7.8 Contour plots of the empirical quantile k_1 (upper), the empirical estimate $\hat{\sigma}_{D,1}$ (middle) and $\hat{\sigma}_{D,3}$ (lower) for the anti-collision case II where one of the wells are shifted in a grid in the NE-plane. The dashed line indicates where $\omega_1 = 1$. 75

7.9 Contour plots of the normal simulated ω_2 (upper), and the relative difference $\frac{\omega_2-\omega_1}{\omega_2}$ (lower) for anti-collision case II. 76

7.10 Anti-collision case III. The offset well is a straight well with $A = 315$ and $I = 60$. The reference well is a straight well with $A = 45$ and $I = 60$. We perform the anti-collision calculations at $D = 3000$ in the offset well. The calculated separation between the wells in this plot is 70.71 meters. 78

7.11 A thorough analysis of the anti-collision case III. Contour plot of ω_1 (upper), ω_3 (middle) and the relative difference $\frac{\omega_3-\omega_1}{\omega_3}$ (lower). 80

7.12 Simulation results of the deviation along the N- (left) and E-coordinate (right) between the expected position from the NIG error model and the calculated position from the minimum curvature formula. The simulation results are made on different measured depths from synthetic straight wells of different inclination angles and azimuth angles. The values along the x-axis are $(D_i - D_s)\sin I_i|\sin A_i|$ (left) and $(D_i - D_s)\sin I_i|\cos A_i|$ (right), where D_s is the measured depth at seabed, and D_i, I_i, A_i are the measured depth, inclination and azimuth of the wells. The green lines are fitted regression models. 81

7.13 Contour plot of ω_1 for anti-collision case I, with the reference well having azimuth angle $A = 60$ instead of $A = 0$. By comparing with Figure 7.3 we see that our results look rotational invariant with respect to the azimuth angle of the reference well. 82

C.1 Three different simulations of the distance D. The upper simulation has large expected value and large variance. The middle simulation has large expected value and small variance. The lower simulation has small expected value and large variance. The normal distribution is a good approximation when D is not too small relative to the uncertainty in D. VI

Notation

Variable	Description
α	Skewness parameter (NIG) or significance level
β	Asymmetry parameter (NIG)
γ	Standardized moment
δ	Magnetic declination angle
δ_{corr}	Correction value of declination angle
η	Parameter vector
ε_i	Error source i
θ	Magnetic dip angle
κ	Dogleg angle
λ	Order of the modified Bessel function, or GH-parameter
Λ	Parameter set (NIG)
μ	Location parameter
μ_k	k'th central moment ($\mu_1 = \mu$)
ξ	Function
π	3.14
ϕ	Density function
φ	Latitude (degrees)
ϱ	Bulk density
ρ	Scale parameter (NIG)
σ	Standard deviation or scale parameter (Normal)
Σ	Covariance matrix
τ	Toolface angle
χ	χ^2-distribution
ω	Separation factor
Ω	Covariance matrix

Variable	Description
A	Azimuth, or constant matrix
A_m	Magnetic azimuth
b	Constant vector
B	Earth's magnetic field strength, field vector
C	Covariance matrix
D	Measured depth, or separation distance
D_0	Calculated minimum centre to centre distance
d	Well diameter
D_V	Vertical depth (TVD)
E	East (coordinate)
EM	Expectation-Maximization algorithm
G	Earth's gravity, or "Global"
GH	Generalised hyperbolic distribution
GIG	Generalised inverse gaussian distribution
h_0	Height above installation
I	Inclination or identity matrix
IG	Inverse gaussian distribution
k	Distribution quantile value
K_λ	Modified Bessel function of the third kind with order λ
m	Central moment estimator
N	North (coordinate), or normal distribution
N_m	Magnetic north
NIG	Normal inverse gaussian distribution
p	Position
P	Position (Random variable)
r	Difference vector
R	Matrix, or "Random"
s^2	Sample variance
S	"Systematic", or sample covariance matrix (S^2)
t	Direction vector
u	Position point
v	Position point
v_k	Vector
V	Vertical (coordinate)
w	Weighting functions
W	Importance sampling weights, or random variable
X	Coordinate axis, or random variable
Y	Coordinate axis, or random variable
Z	Coordinate axis, or random variable

Chapter 1

Introduction

1.1 Introduction

The wells which were drilled in the early ages of the petroleum industry were usually vertical. As the technology has improved, the petroleum industry has been able to drill non-vertical wells. The term ”directional drilling” is used for drilling of non-vertical wells. The need of non-vertical wells is important in order to maximize production from a reservoir. An example of non-vertical wells are offshore wells. A stationary oil platform must be able to drill wells on different locations in order to maximize the production of oil and gas from a reservoir. Both production- and injection-wells often end up with horizontal sections in the reservoir. This leads to wells that have complex geometry.

If all wells were drilled perfectly vertical, it would be impossible for two wells to collide. But since most wells that are drilled today are non-vertical, there will be a risk that a well under construction may collide with an existing well in a reservoir. A collision of two wells is a major safety risk, and would probably cause both environmental- and economical costs, such as damage of equipment and production down-time. If we were able to know the position of the wellbore with absolute certainty, there would be no risk of having a well-collision. In order to avoid well-collisions it is important to have both accurate position calculations, and the most realistic uncertainty estimates as possible. It is equally important to know the position and the uncertainty when considering the probability of hitting the desired target in the reservoir, which is an important factor for economical success.

Measurement while drilling (MWD) is a procedure which is used to compute the position of the wellbore while drilling. MWD-measurements consist of measurements of the Earth’s magnetic field and gravity field for positioning purposes, and various other parameters used in for example geosteering and detection of water and oil. The different measurements are made of different tools, which are mounted behind the wellbore. We focus only on the Earth's gravity- and magnetic field measurements in this thesis. The MWD directional surveying instruments measure orthogonal components of the Earth's magnetic- and gravity field. A sequence of MWD directional measurements in a given part of a well is called an MWD-survey. From now on, when mentioning MWD, it is implicitly meant the MWD directional surveying.

The industry standard of today uses the normal distribution for all error sources concerning MWD-surveys. The normal distribution has the property that it is closed under linearity. The calculation of the wellbore position is done with a set of non-linear equations. These non-linear equations can be linearized, so the position uncertainty is a linear combination of the error uncertainties. Estimates of the uncertainty in the error sources are empirically achieved. Williamson [25] and Ekseth [9] summarize many of the uncertainty estimates which are used as an industry standard today. With the uncertainty estimates of the error sources, under the assumption that the error sources are normally distributed, the position uncertainty also becomes (multivariate) normally distributed. We limit ourselves to use 19 different error sources in this thesis. "The normal error model" is the model where all the 19 error sources are normally distributed, and it represents the model that is currently used as a industry standard.

If a new statistical distribution is used on the error sources in MWD-surveys, we lose the nice properties which follow with the normal distribution. The analytical computation of the position uncertainty is much harder, if not impossible. In this thesis we present methods which allow us to analyze the position uncertainty with non-normal distributions on the error sources. More specific we analyze the use of a Normal Inverse Gaussian (NIG) distribution as an error distribution on the magnetic declination error. There are literature ([22] and [23]) which indicate the need of a heavy-tailed distribution for the declination error. The NIG-distribution is a skewed and heavy tailed statistical distribution which has sufficiently nice properties which make it easy to simulate NIG-distributed realizations and estimate the statistical parameters of the NIG-distribution. The final position uncertainty is not a know statistical distribution, but we are able to numerically evaluate the distribution through an approximation, and we are able to simulate it. The error model with a NIG-distributed declination error and the rest of the error sources normally distributed is denoted the "NIG error model" in the rest of the thesis.

The common practice of anti-collision calculations today is to use the normal error model ([5] and [26]). In the anti-collision setting we denote the planned well as the reference well. All the existing wells around the reference well are offset wells. Anti-collision calculations are done between a point in the reference point, and a candidate point in one of the offset wells. The distance between these points, and the uncertainty of this distance, are important estimates in anti-collision calculations. The common practice is also to approximate the distribution of the distance with a normal distribution. We apply the NIG error model to anti-collision calculations by simulating the distribution of the well-positions, and compare the results against the results we get by using the normal error model. We will use the empirical distribution of the distance between the wells, and not approximate the distance distribution with a normal distribution.

The only computer software used for implementation is MATLAB. Since our purpose is to have comparable results, we need to have full control of the implementation. Real welldata is gathered from other computer software [14] and applied in MATLAB. The computer time of the methods used in this thesis is usually within range of seconds.

1.2 Overview

Chapter 2 gives a summary of MWD magnetic directional surveying. The Earth's magnetic field is introduced, and we discuss models which are used to predict the Earth's magnetic field. We show how measurements of the Earth's magnetic- and gravity field from MWD-surveys are used to compute the position of the wellbore with the minimum curvature method. The methodology in Chapter 2 is established as the industry standard, and this thesis gives no new supplements regarding this theory.

Chapter 3 summarizes the statistical theory which is needed for this thesis. Higher order statistical moments are introduced, and we discuss important properties of the normal distribution and the NIG-distribution. Algorithms for parameter estimation and simulation regarding the NIG-distribution are presented.

Chapter 4 introduces all the error sources which define our error models. The magnitudes of the error sources are used as industry standard, and we use the same magnitudes of the error sources. The normal error model is defined as the error model where all the error sources are normally distributed. The methods for calculating the position uncertainty given that the error sources are normally distributed are discussed in Chapter 4. Chapter 4 is considered as a review of existing theory, as no previous literature have presented it in a statistical setting.

Chapter 5 introduces the use of a NIG-distributed declination error. The NIG error model is introduced as the model where the declination error is NIG-distributed, and the other error sources are normally distributed. We suggest a skewed and heavy-tailed NIG-distribution for the declination error, and we discuss different analytical properties that are preserved under the NIG error model. The parameters of the suggested NIG-distribution are estimated from processed geomagnetic field data from the Tromsø geomagnetic observatory [18]. We compare the position uncertainty of the NIG error model against the normal error model. The contents in Chapter 5 are not found in any previous published literature.

Chapter 6 gives a statistical review of anti-collision calculations, as it is currently used in the oil- and gas-industry. A modification of the existing theory is suggested in order to use the contents of Chapter 5 in anti-collision calculations.

Chapter 7 analyzes the results of this thesis. We analyze three different anticollision cases which are illustrative for extracting the differences between the error models. The main focus of the results are the differences between the normal error model and the NIG error model.

Chapter 8 gives the closing remarks regarding the conclusions and the need of further work on the subject.

Chapter 2

MWD magnetic directional surveying

2.1 Introduction

This chapter gives a theoretical introduction to Measurement While Drilling (MWD) magnetic directional surveying. We show how to compute the position of a wellbore based on measurements from MWD-surveys.

In the early years of the petroleum industry, most wells were drilled vertically to reach the reservoir. The next generation of wells had more advanced geometry in order to hit specific targets in the reservoir. There is often a requirement that the target shall be hit in a specific direction of the well. ”Directional drilling” is a term which is used to describe the drilling of advanced wells. Accurate determination of the wellbore position is important for both safety-, economical- and reliability purposes. There are often numerous wells around a reservoir, and the reservoir targets of the wells are often close to eachother. It is of critical importance to avoid well-collision. A collision with an active well could cause a blowout, which put human lives on risk. The economical costs with a well-collision are also large, since the drillstring most likely has to be pulled out after a collision. It is also important to know the exact location of a well in order to follow the optimal strategy of a production plan.

To determine the position of the wellbore, directional surveys are made during the drilling operations. Directional surveying can be done on a wireline or while drilling (MWD). In MWD-surveys one can measure the gravity with accelerometers and the magnetic field with magnetometers. The accelerometers and magnetometers are equipped in a MWD-tool, which is located behind the wellbore. The Earth’s rotation can be measured on a wireline with gyroscopic instruments, but it is not measured while drilling. Usually both MWD-

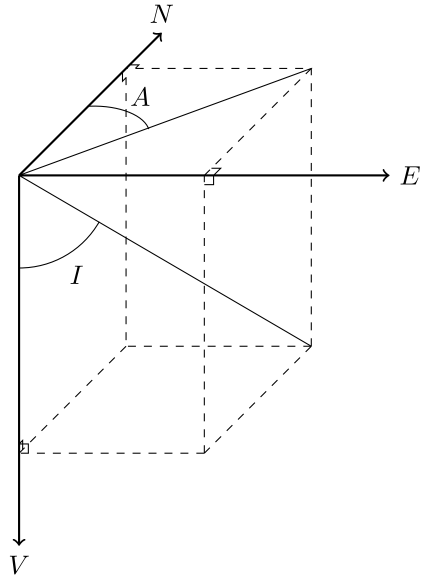


Figure 2.1: Definition of azimuth angle A and the inclination angle I in the NEV coordinate-system. The measured depth D is the "along-hole distance".

surveys and wireline-surveys are done on the same well. But MWD-surveys is the most important survey when considering the risk of having a well-collision, since MWD-surveys are done while the well is drilled. Wireline-surveys are done after the well is drilled, when there is no risk of colliding. We focus only on directional MWD-surveys in this thesis.

Section 2.2 defines the coordinate systems which are used to describe the position and orientation of a wellbore. Section 2.3 introduces the Earth's magnetic field, which is important for directional MWD-surveying. Section 2.4 introduces a model for predicting the Earth's gravity field. Section 2.5 shows how we can calculate the inclination angle I and the azimuth angle A from measurements $G_x, G_y, G_z, B_x, B_y, B_z$ from MWD-surveys. Section 2.6 shows how to calculate the position of a wellbore based on the inclination angle I and the azimuth angle A. We can summarize the framework of Chapter 2 with the following scheme:

$$\begin{bmatrix} G_x \\ G_y \\ G_z \\ B_x \\ B_y \\ B_z \end{bmatrix} \xrightarrow{\text{Section 2.5}} \begin{bmatrix} D \\ I \\ A \end{bmatrix} \xrightarrow{\text{Section 2.6}} \begin{bmatrix} N \\ E \\ V \end{bmatrix} = P$$

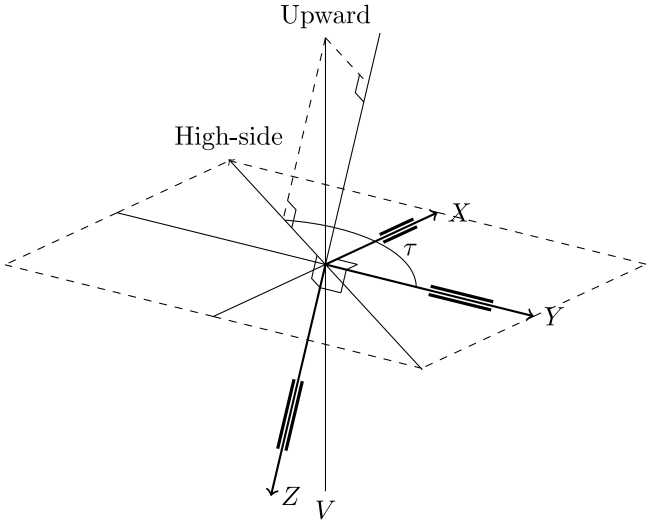


Figure 2.2: Definition of toolface angle τ and the high-side direction in the instrument based coordinate-system, which is defined by the way the sensors are mounted. The z-direction is the down-hole direction, and it is orthogonal to the plane spanned by x and y. The high-side direction is defined as the negative V-direction (upward-direction) perpendicular to the z-axis.

2.2 Coordinate systems

The most common way to describe the position of a wellbore is by using the NEV coordinate-system. NEV is an abbreviation for North-East-Vertical. The NEV coordinate-system allows us to determine a position in latitude, longitude and depth, absolutely or relatively to a platform or a template. To describe the orientation of a wellpath, we introduce two angles: The azimuth angle A and the inclination angle I. The azimuth is the angle between the geographic north and the horizontal projection of the wellpath. The inclination is the angle between vertical axis and the arc of the wellpath. The measured depth D is the "along-hole distance" of the drillstring, which can be interpreted as the length of the arc of the well. Figure 2.1 shows how the angles are defined in the NEV coordinate-system. The magnetic azimuth A_m is also commonly used in MWD-surveys. This angle is defined in the same way as the azimuth, except that A_m is the angle between the magnetic north and the horizontal projection of the wellpath.

We now have sufficient information to describe a wellpath properly. In order to calculate the azimuth A and the inclination I, we need measurements from the MWD-tool. The accelerometer and magnetometer packages in a MWD-tool normally consist of 3 sensors each. They are orthogonally placed, spanning an instrument based coordinate-system. The z-axis in the instrument based coordinate-system defines the down-hole direction, which is the drilling direc-

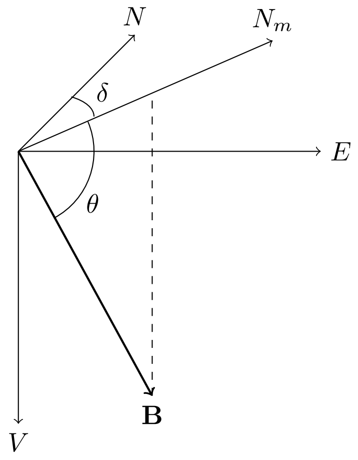


Figure 2.3: Definition of the magnetic field vector in the NEV coordinate-system. The field strength defines the length of B, and the direction of B is defined by the magnetic dip angle θ and the magnetic declination angle δ.

tion. The sensor packages rotate during drilling. In order to keep track of how much the MWD-tool have rotated between each measurement, the high-side direction is defined. The high-side direction is defined as the negative V-direction perpendicular to the z-axis. The angle from the high-side direction to the y-axis is called the toolface angle τ. τ is used to determine the orientation of the sensor packages. Figure 2.2 shows the instrument based coordinate-system and the high-side direction. Note that the toolface angle is undefined for vertical wells, when the z-axis are parallel with the V-axis. An alternative toolface angle, the north toolface τ_n, is used in cases like this. τ_n is the angle between the geographic north and the y-axis. For small inclination angles, we have that $\tau_n \approx A + \tau$.

2.3 Magnetic reference system

The Earth's magnetic field has a central role in directional MWD-surveys. We can describe the Earth's magnetic field, which we denote B, as a vector field. The most common way to describe the vector field is shown in Figure 2.3. All compasses point toward the magnetic north pole, whose direction we denote N_m. The angle between the geographic north and the projection of the Earth's magnetic field vector in the horizontal plane (which is on the line N_m) is called the declination angle δ. The dip angle θ is the angle between the magnetic north and the direction of the magnetic field. The magnetic field strength of B is measured in units Tesla (T).

The Earth's magnetic field can roughly be divided into three different magnetic fields: The main field, the crustal field and the external field. For more details about the magnetic fields, see [16] or [10]. There exists models of how to predict both the main field and the crustal field. Prediction of the Earth's magnetic field is necessary for quality assurance of the MWD-surveys. These models are however only approximations of the reality, so one must expect these models to have uncertainties. The external field is the most unpredictable field, with high-frequent variations of relatively large magnitude. These variations can be observed on stationary geomagnetic observatories [18]. Throughout this thesis we have used ($B = 50000$ nT, $\theta = 75$ and $\delta = 1$) as our reference values for the Earth's magnetic field, which is typical values in the Norwegian sea.

2.4 Gravity reference system

Similar to the magnetic field, we have a model for the Earth's gravity field. There are no common standards in the oil and gas industry how to model the Earth's gravity field. The Earth's gravity field varies much less than the Earth's magnetic field. The presented formula, which is a modification of the International Gravity Formula [24], can be used as reference value for the Earth's gravity field. Let φ be the latitude (in degrees), D_V be the vertical depth (in meters), h_0 be the height of the installation above the mean sea level, and ϱ be the bulk density of the drillstring from the drilling installation to D_V (in g/cm^3). The general formula for predicting the Earth's gravity field is: [17]

$$G = 978030 + 5186 \sin^2 \varphi + 0.309(D_V - h_0) - 0.084 \varrho D_V \tag{2.1}$$

The predicted gravity G is measured in mGal, where 1 mGal = 10^{-5} m/s^2. Equation (2.1) is usually simplified into two formulas:

Offshore formula:

$$G = 978030 + 5186 \sin^2 \varphi + 0.14 D_V$$

Onshore formula:

$$G = 978030 + 5186 \sin^2 \varphi + 0.10 D_V - 0.31 h_0$$

2.5 Connection between the instrument based- and the magnetic reference coordinate system

The NEV coordinate-system and the instrument based coordinate-system are connected through a set of equations. From the accelerometer and magnetometer readings we can calculate the inclination angle I and the magnetic azimuth angle A_m. Measurements in MWD-surveys are typically done approximately every 30 meter. This is because the typical well segment is 30 meters long.

During coupling of two well segments, the sensors have time to transmit the sensor measurements to the installation. When measuring the magnetic field and the gravity field in the x, y, z-directions of the tool coordinate system, we need to have a connection between these measurements and the parameters I and A_m which allow us to compute the relative position change since the last measurement station in the survey. The formulas derived in this section are geometrically derived in [6]. We consider the measurements $G_x, G_y, G_z, B_x, B_y, B_z$. The total magnetic field strength B and the gravity field G are found as the total length of the respective vector:

$$B = \sqrt{B_x^2 + B_y^2 + B_z^2}$$
$$G = \sqrt{G_x^2 + G_y^2 + G_z^2}$$

The accelerometer sensors are directly connected with the inclination angle I of the well, and the toolface angle τ of the MWD-tool. The geometric connection between the accelerometer sensors, the inclination angle I and the toolface angle τ is as follows:

$$G_x = -G \sin I \sin \tau \tag{2.2}$$
$$G_y = -G \sin I \cos \tau \tag{2.3}$$
$$G_z = G \cos I \tag{2.4}$$

The inclination angle I is found from the accelerometer vectors, as G always points in the V-direction:

$$I = \arctan \left(\frac{\sqrt{G_x^2 + G_y^2}}{G_z} \right) \tag{2.5}$$

We see that $I = 0$ if $G = G_z$, which means that the down-hole direction is vertical. The toolface angle τ can be calculated from the accelerometer vectors since the accelerometer sensors are orthogonal to eachother:

$$\tau = \arctan \left(\frac{-G_x}{-G_y} \right) \tag{2.6}$$

As we mentioned in Section 2.2, we see that the toolface angle τ must be modified for vertical wells. This is seen from (2.6) as G_x and G_y are zero for vertical wells. For vertical wells, we use the north toolface $\tau_n \approx A + \tau$. We see that the inclination angle I and the toolface angle τ are fully defined from the accelerometer sensors. It can be shown [9] that the magnetic sensors can be interpreted in the following way:

$$B_x = B\left(\cos\theta \cos I \cos A_m \sin\tau - \sin\theta \sin I \sin\tau + \cos\theta \sin A_m \cos\tau\right) \tag{2.7}$$
$$B_y = B\left(\cos\theta \cos I \cos A_m \cos\tau - \sin\theta \sin I \cos\tau - \cos\theta \sin A_m \sin\tau\right) \tag{2.8}$$
$$B_z = B\left(\cos\theta \sin I \cos A_m + \sin\theta \cos I\right) \tag{2.9}$$

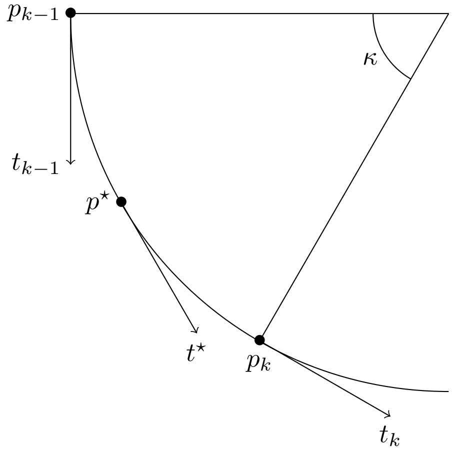


Figure 2.4: Schematic illustration of the notation used in the minimum curvature method. p_{k-1}, $p^\star$ and p_k are positions along the wellpath. MWD-measurements are made in p_{k-1} and p_k, where t_{k-1} and t_k are the direction vectors. $p^\star$ is the position at a desired measured depth $D^\star$, which we must interpolate the position to.

Assuming that we know τ and I from the accelerometer measurements, we can solve (2.7)-(2.9) for the magnetic dip angle θ and the magnetic azimuth A_m:

$$\theta = \arctan\left(\frac{B_z \cos I - [B_x \sin\tau + B_y \cos\tau]\sin I}{\sqrt{([B_x \sin\tau + B_y \cos\tau]\cos I + B_z \sin I)^2 + (B_y \sin\tau - B_x \cos\tau)^2}}\right) \tag{2.10}$$

$$A_m = \arctan\left(\frac{B_x \cos\tau - B_y \sin\tau}{[B_x \sin\tau + B_y \cos\tau]\cos I + B_z \sin I}\right) \tag{2.11}$$

Equations (2.10) and (2.11) are geometrically derived in [6]. The true azimuth angle can be found when we know the declination angle δ of the magnetic field (See Figure 2.3). We can achieve an estimate of the predicted declination angle from the models of the Earth's magnetic field. This estimate, together with the a grid correction which is caused by the meridian convergence, we use a correction value δ_{corr} as an estimate for the declination angle to estimate the true azimuth angle:

$$A = \delta_{corr} + A_m \tag{2.12}$$

For our purposes, δ_{corr} is always given prior to MWD-surveys from magnetic reference values. δ_{corr} is usually the same value for all surveys in a well, even though the model of the Earth's magnetic main field is time- and position-dependent, and the crustal field is position-dependent.

2.6 Minimum curvature methods

When we have computed the inclination I and the azimuth A for a set of new measurements, we are able to compute the relative position change since the last set of measurements. The most commonly used method is the minimum curvature method, and it is thoroughly discussed in [20]. The minimum curvature method has emerged as the accepted industry standard for calculating the well-position. Figure 2.4 shows an illustration of the notation used in the minimum curvature method. Let p_{k-1} be the position of the previous measurements (In NEV-coordinate system). New MWD-measurements are done in position p_k, which is currently unknown. From the measurements in p_{k-1} and the new measurements, we are able to compute the new position p_k. Let I_{k-1} and I_k denote the two inclination angles derived from the measurements in p_{k-1} and p_k. Let A_{k-1} and A_k denote the respective azimuth angles. A direction vector t_{k-1} can be calculated from the inclination I_{k-1} and the azimuth A_{k-1}:

$$t_{k-1} = \begin{bmatrix} \sin I_{k-1} \cos A_{k-1} \\ \sin I_{k-1} \sin A_{k-1} \\ \cos I_{k-1} \end{bmatrix}, \tag{2.13}$$

and likewise for t_k. Let $\Delta D = D_k - D_{k-1}$ be the difference in measured depth. We define κ to be the measure of the change in inclination angle in a well, which is called the dogleg angle. An expression for the dogleg angle can be computed [20]:

$$\kappa = 2 \arcsin \left\{ \sqrt{\sin^2 \left(\frac{I_k - I_{k-1}}{2}\right) + \sin I_{k-1} \sin I_k \sin^2 \left(\frac{A_k - A_{k-1}}{2}\right)} \right\} \tag{2.14}$$

We see from Figure 2.4 and (2.14) that κ is zero if the well has constant inclination and azimuth. Thorogood and Sawaryn [20] derives the following formula for computing p_k:

$$p_k = p_{k-1} + \frac{\Delta D f(\kappa)}{2} \begin{bmatrix} \sin I_{k-1} \cos A_{k-1} + \sin I_k \cos A_k \\ \sin I_{k-1} \sin A_{k-1} + \sin I_k \sin A_k \\ \cos I_{k-1} + \cos I_k \end{bmatrix}, \tag{2.15}$$

where $f(\kappa) = \frac{\tan \frac{\kappa}{2}}{\frac{\kappa}{2}}$. Another minimum curvature formula which is important for our purposes is to calculate the position $p^\star$ on a given measured depth $D^\star$. This is a minimum curvature interpolation method which is important if we want to compute the position of a well on a equidistant grid in measured depth. Let p_{k-1} and p_k be two calculated positions along the wellpath. Let $\Delta D^\star = D^\star - D_{k-1}$ be the distance between p_{k-1} and $p^\star$, and let $\kappa^\star$ be the dogleg angle between p_k and $p^\star$. The direction vector $t^\star$ can be calculated from the following formula [20]:

$$t^\star = \frac{\sin\left([1 - \frac{\Delta D^\star}{\Delta D}]\kappa\right)}{\sin \kappa} t_{k-1} + \frac{\sin\left(\frac{\Delta D^\star}{\Delta D}\kappa\right)}{\sin \kappa} t_k$$

In a similar manner as in (2.15), we can now compute the position vector $p^\star$ at the desired measured depth:

$$p^\star = p_{k-1} + \frac{\Delta D^\star f(\kappa \frac{\Delta D^\star}{\Delta D})}{2}(t_{k-1} + t^\star)$$

Chapter 3

Statistical distributions

3.1 Introduction

This chapter summarizes the statistical distributions which are used in this thesis. The properties of the statistical distributions which are used in the thesis are also shown. Section 3.2 explains the theory behind statistical moments, which are important for this thesis. Section 3.3 introduces the normal distribution. Section 3.4 introduces the NIG-distribution, and algorithms for estimating the NIG-parameters and simulation of NIG-distributed random variables are presented. Section 3.5 reflects over other possible statistical distributions than the NIG-distribution which could have been applied in this thesis.

3.2 Statistical moments

The statistical moments are an important class of expectations of a distribution [7]. Let X be a random variable with density function $f_X(x)$, expectation μ and variance σ^2. The k'th central moment, which is the k'th moment about the mean μ, is defined as:

$$\mu_k = \int_{-\infty}^{\infty} (x-\mu)^k f_X(x) \mathrm{d}x \quad ; \ k = 1, 2, ... \tag{3.1}$$

We see from (3.1) that the first central moment $\mu_1 = 0$. For $k = 2$, we recognize (3.1) as the definition of the variance σ^2. The higher order central moments are used to describe other properties of a statistical distribution. We define the k'th standardized moment γ_k:

$$\gamma_k = \frac{\mu_k}{\sigma^k} \quad ; \ k = 1, 2, ... \tag{3.2}$$

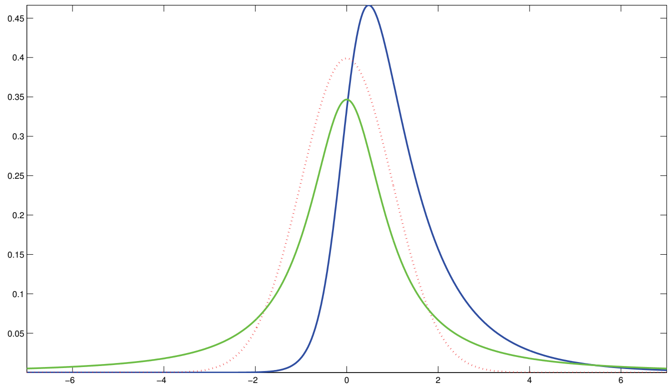


Figure 3.1: Three different probability density functions. The dashed red line is a standard normal distribution, the green line is a symmetric distribution with heavy tails, and the blue line is a right-skewed distribution.

We see from (3.2) that the first standardized moment $\gamma_1 = 0$ since $\mu_1 = 0$. For $k = 2$, the standardized moment $\gamma_2 = 1$. The third and fourth standardized moments are called the skewness and kurtosis, respectively. The skewness and kurtosis are important properties of a statistical distribution. If the skewness $\gamma_3 > 0$, the distribution is called positive-skew, or right-skewed. The blue distribution in Figure 3.1 is an example of a right-skewed distribution. A symmetric distribution, such as the normal distribution, has $\gamma_3 = 0$. The kurtosis γ_4 is a measure of the peakedness or the flatness of a distribution [7]. A high kurtosis results in a sharp peak and heavy tails of a distribution. The green distribution in Figure 3.1 is an example of a distribution with $\gamma_4 = 30$. The normal distribution (the dashed red line in Figure 3.1) has $\gamma_4 = 3$, and a kurtosis larger than this results in heavier tails than the normal distribution. We can estimate the k'th central moment from a sample $X_1, .., X_n$ with the sample estimator m_k of μ_k:

$$m_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k \quad k = 1, 2, .. \tag{3.3}$$

From (3.3) follows the estimators of sample skewness $\hat{\gamma}_3$ and sample kurtosis $\hat{\gamma}_4$:

$$\hat{\gamma}_3 = \frac{m_3}{m_2^{3/2}} \tag{3.4}$$

$$\hat{\gamma}_4 = \frac{m_4}{m_2^2} \tag{3.5}$$

3.3 The normal distribution

The normal distribution is by far the most widely used statistical distribution. It has a wide range of applications, and is a limit distribution in many cases. A random variable X is normally distributed with expectation μ and variance σ^2 if X has the following density function:

$$f_X(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad , \; x \in \mathbb{R} \tag{3.6}$$

A commonly used abbreviation for "X is normally distributed with expectation μ and variance σ^2" is $X \sim N(\mu, \sigma^2)$. The univariate normal distribution is a member of the location-scale family of univariate probability distributions. If $Z \sim N(0,1)$, we have the property that $X = \mu + \sigma Z$. From this property, μ is often called the location parameter, and σ is often called the scale parameter.

If we have n independent random variables $X_1, .., X_n$, which we assume are normally distributed with expectation μ and variance σ^2, we can use the maximum likelihood estimators (MLE) of μ and σ^2 to fit a suitable normal distribution. These estimators are available in every basic statistical software, and are as follows:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{3.7}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \hat{\mu})^2 \tag{3.8}$$

Sometimes the maximum likelihood estimator $\hat{\sigma}^2$ (3.8) is replaced by the sample variance estimator $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$, if it is preferable to have an unbiased estimator of σ^2. The normal distribution is closed under linearity. If

$$Y = \sum_{i=1}^{n} a_i X_i + b \quad ,$$

where $X_i \sim N(\mu_i, \sigma_i^2)$ and the X_i's are uncorrelated, we have that

$$Y \sim N\left(b + \sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

The normal distribution can be extended to a multivariate distribution. Let $X = [X_1, .., X_k]^T$ be a random vector with k elements. If the elements of X are normally distributed with known expectation and variance, and the covariance structure between the elements is known, we can describe the statistical distribution of X with a multivariate normal distribution. The random vector X of dimension k is multivariate normal with parameter vector μ and covariance matrix Σ, X has the following density function:

$$\phi_k(x;\mu,\Sigma) = \frac{1}{(2\pi)^{k/2}}|\Sigma^{-1/2}| \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\} \quad , \; x \in \mathbb{R}^k \tag{3.9}$$

We use the abbreviation $X \sim N_k(\mu, \Sigma)$ if X is multivariate normally distributed. The multivariate normal distribution is a member of the location-scale family of multivariate distributions. If $Z \sim N_k(0, I)$, where I is the identity matrix, we have the property that $X = \mu + \Sigma^{1/2} Z$. $\Sigma^{1/2}$ is the Cholesky-decomposition of Σ. If we have n independent random vectors $X_1, .., X_n$, which we assume are multivariate normally distributed with expectation μ and covariance matrix Σ, we have the following maximum likelihood estimators:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{3.10}$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{\mu})(X_i - \hat{\mu})^T \tag{3.11}$$

As in the univariate case, the maximum likelihood estimator $\hat{\Sigma}$ (3.11) can be replaced by the sample variance estimator $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \hat{\mu})(X_i - \hat{\mu})^T$, if it is preferable to have an unbiased estimator of Σ. As for the univariate case, the multivariate normal distribution is also closed under linearity. If

$$Y = \sum_{i=1}^{n} A_i X_i + b \quad ,$$

where $X_i \sim N_k(\mu_i, \Sigma_i)$, $A_i \in \mathbb{R}^{m \times k}$ and $b \in \mathbb{R}^m$, we have that

$$Y \sim N_m \left(b + \sum_{i=1}^{n} A_i \mu_i, \sum_{i=1}^{n} A_i \Sigma_i A_i^T \right)$$

3.4 The normal inverse gaussian distribution

3.4.1 Density function

The normal inverse gaussian (NIG) distribution is a distribution which can describe data with skewness and fat tails. Most of the literature about the NIG distribution can be found in [1], [13] or [11]. It is a special case of the generalised hyperbolic (GH) distribution. The parameters of a univariate NIG are $\Lambda = (\mu, \rho, \alpha, \beta)$. The density function of $X \sim NIG(\mu, \rho, \alpha, \beta)$ is:

$$f_X(x) = \frac{\alpha \rho K_1 \left(\alpha \sqrt{\rho^2 + (x-\mu)^2} \right)}{\pi \sqrt{\rho^2 + (x-\mu)^2}} \exp \left\{ \rho \sqrt{\alpha^2 - \beta^2} + \beta (x - \mu) \right\} \quad , \; x \in \mathbb{R} \tag{3.12}$$

We have the restrictions $\mu \in \mathbb{R}$, $\alpha > 0$, $\rho > 0$ and $0 < |\beta| \leq \alpha$ on the parameters of this distribution. For easier parametrization, one can use $\gamma = \sqrt{\alpha^2 - \beta^2}$, and let $\gamma > 0$. Both γ and α are used in this thesis, although it is important to remember that the NIG-distribution only has 4 parameters. $K_1(\cdot)$ is the modified Bessel function of the third kind and of order 1. The modified Bessel function of the third kind with order λ is the following function:

$$K_\lambda(x) = \frac{1}{2}\int_0^\infty w^{\lambda-1} \exp\left\{-\frac{1}{2}(xw - \frac{x}{w})\right\} \mathrm{d}w$$

The NIG distribution is a member of the location family of univariate distributions. If $Z \sim NIG(0, \rho, \alpha, \beta)$, we have that $X = \mu + Z \sim NIG(\mu, \rho, \alpha, \beta)$. μ is often called the location parameter, also for the NIG distribution. The mean, variance, skewness and kurtosis of a NIG distribution can be found by the following formulas:

$$E[X] = \mu + \frac{\rho\beta}{\gamma} \tag{3.13}$$

$$\text{Var}[X] = \frac{\rho\alpha^2}{\gamma^3} \tag{3.14}$$

$$\text{Skew}[X] = \frac{3\beta}{\alpha\sqrt{\rho\gamma}} \tag{3.15}$$

$$\text{Kurt}[X] = 3\left(1 + 4\frac{\beta^2}{\alpha^2}\right)\frac{1}{\rho\gamma} \tag{3.16}$$

ρ is often called a scale parameter, even though the NIG distribution is not a member of the location-scale family. We see that ρ occurs in all the expressions (3.13)-(3.16). If $\beta = 0$, we see that the skewness (3.15) becomes 0. Of this reason, β is often called the asymmetry parameter. If $\beta > 0$, the distribution is right-skewed, and if $\beta < 0$, the distribution is left-skewed. α and β occur both separately, and together in γ, in (3.13)-(3.16). Together they determine the heaviness of the tails. The heaviest tail decays as:

$$f_X(x) \propto |x|^{-3/2} \exp\left\{-\alpha|x| + |\beta x|\right\} \begin{cases} \beta < 0 & \text{and} \quad x \to -\infty \\ \beta > 0 & \text{and} \quad x \to +\infty \end{cases}$$

The lightest tail decays as:

$$f_X(x) \propto |x|^{-3/2} \exp\left\{-\alpha|x| - |\beta x|\right\} \begin{cases} \beta < 0 & \text{and} \quad x \to +\infty \\ \beta > 0 & \text{and} \quad x \to -\infty \end{cases}$$

The normal distribution is closed under linearity. The NIG distribution is only closed under linearity for special cases of linear transformations. If $Y = \sum_{i=1}^{n} X_i$, where $X_i \sim NIG(\mu_i, \rho_i, \alpha, \beta)$, we have that

$$Y \sim NIG\left(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\rho_i, \alpha, \beta\right)$$

If $Y = aX + b$, where a and b are scalars, and $X \sim NIG(\mu, \rho, \alpha, \beta)$, we have that

$$Y \sim NIG\left(a\mu + b, |a|\rho, \frac{\alpha}{|a|}, \frac{\beta}{a}\right)$$

The proofs of these properties are shown in Appendix A, as we failed to find them in other literature.

The density function (3.12) involves the modified Bessel function, which complicates all calculations regarding X. We can avoid the modified Bessel function by introducing a new random variable. Suppose we have a random variable Z which is inverse gaussian distributed, $Z \sim IG(\rho, \gamma)$, i.e.:

$$f_Z(z) = \frac{1}{\sqrt{2\pi z^3}} \rho \exp\left\{\rho\gamma - \frac{1}{2}(\frac{\rho^2}{z} + \gamma^2 z)\right\} \tag{3.17}$$

Let the conditional distribution $X|Z$ be normal distributed, $X|Z = z \sim N(\mu + \beta z, z)$. The joint density of (X, Z) is:

$$\begin{aligned} f(x,z) &= f(x|z) \times f(z) \\ &\propto \frac{\rho}{z^2} \exp\left\{\rho\gamma - \beta(\mu - x) + \mu\frac{x}{z} - \frac{1}{2}(\beta^2 + \gamma^2)z - \frac{1}{2}(\mu^2 + \rho^2)z^{-1}\right\} \end{aligned} \tag{3.18}$$

If we compute the marginal distribution of X from (3.18), we get that X is marginally NIG-distributed, $X \sim NIG(\mu, \rho, \alpha, \beta)$. The proof is shown in Appendix B. We can observe that the joint density (3.18) is also an exponential family:

$$f(x,z) = g(\Lambda) h(x,z) \exp\left\{\sum_{k=1}^{4} \xi_k(\Lambda) t_k(x,z)\right\} \tag{3.19}$$

Here $g(\Lambda) = \rho e^{\rho\gamma - \beta\mu}$, $t_1(x,z) = x$, $t_2(x,z) = \frac{x}{z}$, $t_3(x,z) = \frac{z}{2}$, $t_4(x,z) = \frac{1}{2z}$, and the $\xi_i(\cdot)$'s are:

$$\begin{aligned} \xi_1(\Lambda) &= \beta \\ \xi_2(\Lambda) &= \mu \\ \xi_3(\Lambda) &= -(\beta^2 + \gamma^2) \\ \xi_4(\Lambda) &= -(\mu^2 + \rho^2) \end{aligned}$$

The sufficient statistics for (X, Z) can be easily obtained by (3.19). They are given by $t_k(X, Z)$ $(k = 1, 2, 3, 4)$. We can use the property that the joint density of (X, Z) is an exponential family when we estimate the parameters by the EM-algorithm.

3.4.2 Parameter estimation by EM-algorithm

We must use numerical, iterative methods to estimate the parameters in this distribution. One of the most reasonable methods to use is the Expectation-Maximization-algorithm (EM-algorithm). Details of the EM-algorithm in general can be found in [7] and [8]. The specific EM-algorithm applied to the NIG-distribution is found in [1]. The EM-algorithm converges to the MLE, and is suitable for situations where we have missing data. We use the property that $X|Z = z \sim N(\mu + \beta z, z)$, and consider Z as our missing data.

E-step In the E-step, we compute the expected value of the sufficient statistics for the missing data conditioned on the observed data. From the joint density (3.18), it can be shown that $Z|X \sim GIG(-1, \sqrt{\rho^2 + (x-\mu)^2}, \alpha)$, where GIG is short for the generalised inverse gaussian distribution. Of the joint density (3.19) being an exponential family, and $f(z|x) \propto f(x,z)$, we see that the sufficient statistics of a GIG-distribution are Z and Z^{-1}. The moments of a $GIG(\lambda, a, b)$ are given by [12]:

$$E[Z^r] = \left(\frac{a}{b}\right)^r \frac{K_{\lambda+r}(ab)}{K_\lambda(ab)}$$

Our GIG-distribution has $\lambda = -1$. Hence, we set Z_i and Z_i^{-1} equal to their expected value in the E-step, which are:

$$E[Z_i|X_i = x_i] = \frac{\sqrt{\rho^2 + (x_i-\mu)^2}}{\alpha} \frac{K_0(\alpha\sqrt{\rho^2 + (x_i-\mu)^2})}{K_1(\alpha\sqrt{\rho^2 + (x_i-\mu)^2})} \quad ; \; i = 1, .., n$$
$$E[Z_i^{-1}|X_i = x_i] = \frac{\alpha}{\sqrt{\rho^2 + (x_i-\mu)^2}} \frac{K_2(\alpha\sqrt{\rho^2 + (x_i-\mu)^2})}{K_1(\alpha\sqrt{\rho^2 + (x_i-\mu)^2})} \quad ; \; i = 1, .., n$$

M-step In the M-step, we know the current values of the missing data from the E-step, and can compute estimates for the parameters. From the property that $f(x, z; \mu, \rho, \alpha, \beta) = f(x|z; \mu, \beta) f(z; \rho, \alpha, \beta)$ the likelihood can be split into two separate likelihoods. From $X|Z = z \sim N(\mu + \beta z, z)$ we can easily compute $\beta^{(k+1)}$ and $\mu^{(k+1)}$. This likelihood is:

$$\begin{aligned} L_1(\mu, \beta|\mathbf{x}, \mathbf{z}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi z_i}} \exp\left\{-\frac{1}{2z_i}(x_i - \mu - \beta z_i)^2\right\} \\ &\propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} \frac{1}{z_i}(x_i - \mu - \beta z_i)^2\right\} \end{aligned}$$

Observe that finding estimates for μ and β is equal to solving a regression problem. This means that if we have a sample $X_1, .., X_n$, all X_i's have different variances Z_i. We differentiate the log-likelihood $\log L_1(\mu, \beta)$ with respect to μ and β and set to zero:

$$\frac{\partial \log L_1(\mu, \beta)}{\partial \mu} = \sum_{i=1}^{n} \frac{1}{z_i}(x_i - \mu - \beta z_i) = 0 \tag{3.20}$$
$$\frac{\partial \log L_1(\mu, \beta)}{\partial \beta} = \sum_{i=1}^{n} (x_i - \mu - \beta z_i) = 0 \tag{3.21}$$

By solving (3.20) and (3.21) for β and μ we get the updates $\beta^{(k+1)}$ and $\mu^{(k+1)}$:

$$\beta^{(k+1)} = \frac{\sum_{i=1}^{n} \frac{x_i}{z_i} - \bar{x}\sum_{i=1}^{n} \frac{1}{z_i}}{n - \bar{z}\sum_{i=1}^{n} \frac{1}{z_i}} \tag{3.22}$$
$$\mu^{(k+1)} = \bar{x} - \beta^{(k+1)}\bar{z} \tag{3.23}$$

For computation of $\rho^{(k+1)}$ and $\gamma^{(k+1)}$, we use that $Z \sim IG(\rho, \gamma)$. The likelihood is:

$$\begin{aligned} L_2(\rho, \gamma | \mathbf{x}, \mathbf{z}) &= \prod_{i=1}^{n} \frac{\rho}{\sqrt{2\pi z^3}} \exp\left\{\rho\gamma - \frac{1}{2}(\frac{\rho^2}{z_i} + \gamma^2 z_i)\right\} \\ &\propto \rho^n \exp\left\{n\rho\gamma - \frac{1}{2}(\rho^2 \sum_{i=1}^{n} \frac{1}{z_i} + \gamma^2 \sum_{i=1}^{n} z_i)\right\} \end{aligned}$$

We differentiate the log-likelihood $\log L_2(\rho, \gamma)$ with respect to ρ and γ, and set to zero:

$$\frac{\partial \log L_2(\rho, \gamma)}{\partial \rho} = \frac{n}{\rho} + n\gamma - \rho \sum_{i=1}^{n} \frac{1}{z_i} = 0 \tag{3.24}$$

$$\frac{\partial \log L_2(\rho, \gamma)}{\partial \gamma} = n\rho - \gamma \sum_{i=1}^{n} z_i = 0 \tag{3.25}$$

By solving (3.24) and (3.25) for ρ and γ we get the updates $\rho^{(k+1)}$ and $\gamma^{(k+1)}$:

$$\rho^{(k+1)} = \sqrt{\frac{n}{\sum_{i=1}^{n} \frac{1}{z_i} - \frac{n}{\bar{z}}}} \tag{3.26}$$

$$\gamma^{(k+1)} = \frac{\rho^{(k+1)}}{\bar{z}} \tag{3.27}$$

The update for α are given by the variable transformation:

$$\alpha^{(k+1)} = \sqrt{(\gamma^{(k+1)})^2 + (\beta^{(k+1)})^2} \tag{3.28}$$

By doing the EM-algorithm iteratively, setting the sufficient statistics Z_i and Z_i^{-1} equal to their expected values, and updating the parameters, the parameter updates $(\mu^{(k+1)}, \rho^{(k+1)}, \alpha^{(k+1)}, \beta^{(k+1)})$ converge to the maximum likelihood estimates $(\hat{\mu}, \hat{\rho}, \hat{\alpha}, \hat{\beta})$. Suitable starting values for the parameters are the moment estimators. Let m_l $(l = 2, 3, ..)$ be the sample moment defined in (3.3), and the sample skewness and the sample kurtosis be defined in (3.4) and (3.5). By setting the sample mean, sample variance, sample skewness and sample kurtosis equal to the theoretical values for the mean, variance, skewness and kurtosis in (3.13)-(3.16), and solve for the parameters, we get the following moment estimators as initial values in the EM-algorithm:

$$\begin{aligned} \beta^{(0)} &= \frac{m_3 m_2 \psi^2}{3} \\ \alpha^{(0)} &= \sqrt{\psi^2 + (\beta^{(0)})^2} \\ \rho^{(0)} &= \frac{m_2^2 \psi^3}{(\beta^{(0)})^2 + \psi^2} \\ \mu^{(0)} &= m_1 - \frac{\beta^{(0)} \rho^{(0)}}{\psi} \end{aligned}$$

$$\psi = \frac{3}{m_2\sqrt{3m_4 - 5m_3^2}}$$

The experiences with the EM-algorithm is mixed. The convergence can be very slow in some cases, and in other cases the convergence is fast. In most cases it is sufficient to have a relatively high tolerance limit, because the estimates after a few iterations are often sufficiently good. The likelihood function seems to have very low variability around the maximum, so the estimates can keep changing for a very long time if the tolerance limit is set too small. In some cases, the moment estimators are even a sufficiently good parameter choice.

3.4.3 Simulation

Simulation of independent NIG-distributed random variables can be done very efficiently. In [2], the following algorithm is presented for simulating $X \sim NIG(\mu, \rho, \alpha, \beta)$:

- Sample $Z \sim IG(\rho, \gamma)$
- Sample $Y \sim N(0, 1)$
- Set $X = \mu + \beta Z + \sqrt{Z}Y$

The following algorithm can be used for simulating $Z \sim IG(\rho, \gamma)$ if no available statistical software have this function built in:

- Sample $V \sim \chi_1^2$
- Set $\tau = \frac{\rho}{\gamma}$ and $\lambda = \rho^2$
- Compute $Z_1 = \tau + \frac{\tau}{2\lambda}\left(\tau V - \sqrt{4\tau V\lambda + (\tau V)^2}\right)$ and $Z_2 = \frac{\tau^2}{Z_1}$
- Compute $p = \frac{\tau}{\tau + Z_1}$
- Set $Z = \begin{cases} Z_1 & \text{with probability} \quad p \\ Z_2 & \text{with probability} \quad p \end{cases}$

3.5 Other skewed and heavy-tailed distributions

We have focused on the normal inverse gaussian distribution in this thesis. There are however many alternative distributions that can be used instead of the NIG-distribution. This section presents other alternative distributions that can be used to model skewness and kurtosis in different ways, and justifies why we have chosen to use the NIG-distribution in this thesis.

The NIG-distribution is a special case of the generalised hyperbolic (GH) distribution [2], which has the following density function:

$$f_X(x) = \frac{(\alpha^2-\beta^2)^{\lambda-1/2} K_{\lambda-1/2}\left(\alpha\sqrt{\rho^2+(x-\mu)^2}\right)\exp\{\beta(x-\mu)\}}{\sqrt{2\pi}\alpha^{\lambda-1/2}\rho^{\lambda}K_{\lambda}\left(\rho\sqrt{\alpha^2-\beta^2}\right)\left(\sqrt{\rho^2-(x-\mu)^2}\right)^{1/2-\lambda}} \quad ; \; x \in \mathbb{R}$$

$K_j(\cdot)$ is the modified Bessel function of third kind with order j. We see that the NIG-distribution is a GH-distribution with $\lambda = -\frac{1}{2}$. Since the NIG-distribution is a special case of the GH-distribution, we could have studied the GH-distribution in detail in this thesis. Many of the properties for the NIG-distribution shown in Section 3.4 can be generalised for the GH-distribution. But the expressions for $E(X)$ and $\mathrm{Var}(X)$ include the modified Bessel function, and the parameter estimation becomes more complicated. Since we are interested in a statistical distribution which flexibly can be set with respect to the four first moments, we only need four parameters to uniquely determine the distribution. The NIG-distribution is considered to be sufficiently flexible for our purposes.

If we are not interested in a distribution which models the kurtosis in a flexible way, the skew-normal distribution (SN) [3] can be used. The SN-distribution is a non-symmetric normal distribution, which has the following density function:

$$f_X(x;\mu,\sigma,\alpha) = 2\phi(x;\mu,\sigma)\Phi(\alpha(x-\mu)) \quad , \; x \in \mathbb{R}$$

$\phi(\cdot)$ is the density function of a normal distribution (3.6), and $\Phi(\cdot)$ is the distribution function of a standard normal distribution. α is often called the skewness parameter. This distribution is not considered to be flexible enough for our purposes, since the parameters μ, σ and α lack to model the kurtosis in a flexible way. But it can be useful for cases where data are non-symmetric and the tails decay according to a normal distribution.

In statistical literature there are two different definitions of a Skew-t distribution. We can introduce a Skew-t distribution (ST) in the same way as for the SN-distribution if we want to model heavier tails. For more details regarding this distribution, see [4] or [10]. This ST-distribution involves a parameter regarding the degrees of freedom, which is hard to interpret and harder to estimate. Another Skew-t distribution is introduced in [2] as a special case of the GH-distribution. A Skew-t distribution can be used for modelling skewness and tails that are slightly heavier than in the normal distribution.

Chapter 4

Normally distributed position uncertainty in MWD magnetic directional surveying

4.1 Introduction

This chapter gives an introduction to the statistics which is commonly used in the oil and gas industry to describe the position uncertainty in directional drilling.

The MWD magnetic directional surveys that are made during drilling are basically measurements of components of the Earth's magnetic- and gravity field. The wellbore position are calculated with a set of non-linear equations, as shown in Chapter 2. But the MWD-measurements are physical measurements which are affected by several error sources. These error sources propagate through the non-linear equations and cause an uncertainty in the wellbore position. Accurate determination of the wellbore position is important for both safety-, economical- and reliability purposes. Because of this, it is of critical importance to have a good overview of the error sources that contribute significantly to the position uncertainty, and it is also important to know how these error sources propagate through MWD-surveys.

Most of the theory from this chapter is derived from Ekseth [9]. Chapter 2 explained how to compute the position of a wellbore based on measurements from MWD-surveys. In this chapter we introduce different error sources which cause significant effect on the uncertainty of the wellbore position. These error sources define our error model for directional MWD-surveying. Section 4.2 explains the error sources which we include in our error model. We limit the number of error sources to the most significant ones. We let $\varepsilon_1, .., \varepsilon_{19}$ denote the error sources which we include in this thesis. The weighting functions w, which explains how an error affect the measured depth D, inclination angle I and the azimuth angle A, are also introduced in Section 4.2. Section 4.3 explains how to compute the position uncertainty assuming that all error sources are normally distributed. The assumption of normally distributed error sources has become an established industry standard, as the resulting position uncertainty is normally distributed. This theory is mainly based on linearizations of the minimum curvature formula (2.15), and these linearizations are valid also for error sources with other statistical distributions than the normal distribution. We can summarize the framework of Chapter 4 with the following scheme:

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{19} \end{bmatrix} \xrightarrow[w]{\text{Section 4.2}} \begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix} \xrightarrow[(2.15)]{\text{Section 4.3}} \begin{bmatrix} \mathrm{d}N \\ \mathrm{d}E \\ \mathrm{d}V \end{bmatrix} = \mathrm{d}P$$

4.2 Error model

There are numerous sources of error which contribute to the position uncertainty. Some error sources cause larger uncertainties than other. The chosen error model must be realistic, and include the most important error sources. The more error sources that are included in an error model, the more complicated the model becomes. An important assumption is that all the error sources are independent of eachother. This section defines the error model which we use in this thesis. The different error sources are also explained in this section. The chosen error model includes the most important error sources that contribute to the position uncertainty, according to Williamson [25] and Brooks & Wilson [5]. Table 4.1 gives an overview of the error sources which are included in the error model. The standard deviations that are displayed are suggested values from Williamson [25]. In this section we also introduce weighting functions, which show how the error sources cause errors in measured depth D, inclination angle I and azimuth angle A.

4.2.1 Accelerometer bias errors

The accelerometer package consists of 3 sensors that measure components of the Earth's gravity field. The sensors are mechanical devices, and need to be recalibrated at regular intervals to secure that the accuracy of the measurements is within given limits. It is natural to believe that the sensor errors are systematic within each MWD-survey, but random between different MWD-surveys. The accelerometer bias error cause the sensor readings to be greater or smaller than the true value.

Error number	Description	Standard deviation	Propagation mode
1	X-Accelerometer bias	0.0039 m/s^2	S
2	Y-Accelerometer bias	0.0039 m/s^2	S
3	Z-Accelerometer bias	0.0039 m/s^2	S
4	X-Accelerometer scale	0.0005	S
5	Y-Accelerometer scale	0.0005	S
6	Z-Accelerometer scale	0.0005	S
7	X-Magnetometer bias	70 nT	S
8	Y-Magnetometer bias	70 nT	S
9	Z-Magnetometer bias	70 nT	S
10	X-Magnetometer scale	0.0016	S
11	Y-Magnetometer scale	0.0016	S
12	Z-Magnetometer scale	0.0016	S
13	Depth reference	0.35 m	R
14	Depth scale factor	$6 \cdot 10^{-4}$	S
15	Depth stretch type	$2.5 \cdot 10^{-7}$ m^{-1}	G
16	Declination	0.36°	G
17	Declination B_H-dependent	5000° nT	G
18	Sag	0.2°	S
19	Axial magnetisation in drillstring	150 nT	S

Table 4.1: Overview of the different error sources used in this thesis. All the error sources in this table are considered to be uncorrelated. The way these error sources affect MWD-surveys is shown in the propagation mode column. (R: Random, S: Systematic, G: Global). The standard deviations in the table are from [25].

We assume that the 3 sensors have 3 independent biases. Let ε_1, ε_2 and ε_3 denote the accelerometer bias errors, following the numbering from Table 4.1. The errors ε_1, ε_2 and ε_3 are independent random variables. A bias error ε_1 in the X-accelerometer causes an error $\mathrm{d}G_x = \varepsilon_1$ in the sensor reading. The accelerometer errors caused by the accelerometer bias errors are:

$$
\begin{aligned}
\mathrm{d}G_x &= \varepsilon_1 \\
\mathrm{d}G_y &= \varepsilon_2 \\
\mathrm{d}G_z &= \varepsilon_3
\end{aligned}
$$

The accelerometer measurements are used both in computation of inclination I and azimuth A. An error $\mathrm{d}G_x$ in the X-accelerometer causes the errors $\mathrm{d}I$ and $\mathrm{d}A$ in I and A. If we assume that the errors are small, we can approximate $\mathrm{d}I$

and dA from (2.5) and (2.11):

$$\mathrm{d}I \approx \frac{\partial I}{\partial G_x}\mathrm{d}G_x = \frac{\partial I}{\partial G_x}\varepsilon_1 \tag{4.1}$$

$$\mathrm{d}A \approx \frac{\partial A}{\partial I}\frac{\partial I}{\partial G_x}\mathrm{d}G_x = \frac{\partial A}{\partial I}\frac{\partial I}{\partial G_x}\varepsilon_1 \tag{4.2}$$

The same can be done for the bias errors in the two other accelerometers. We define a weighting function $w_1 = [\frac{\partial D}{\partial G_x}, \frac{\partial I}{\partial G_x}, \frac{\partial A}{\partial G_x}]^T$ for the X-accelerometer bias, and likewise for the two other accelerometers. It can be shown that the weighting functions for the accelerometer bias errors are as follows:

$$w_1 = \frac{1}{G}\begin{bmatrix} 0 \\ -\cos I \sin\tau \\ (\cos I \sin A_m \sin\tau - \cos A_m \cos\tau)\tan\theta + \cot I \cos\tau \end{bmatrix} \tag{4.3}$$

$$w_2 = \frac{1}{G}\begin{bmatrix} 0 \\ -\cos I \cos\tau \\ (\cos I \sin A_m \cos\tau + \cos A_m \sin\tau)\tan\theta - \cot I \sin\tau \end{bmatrix} \tag{4.4}$$

$$w_3 = \frac{1}{G}\begin{bmatrix} 0 \\ -\sin I \\ \tan\theta \sin I \sin A_m \end{bmatrix} \tag{4.5}$$

4.2.2 Accelerometer scale errors

The accelerometer scale errors are justified in the same way as the accelerometer bias errors. The scale errors are also assumed to be systematic within each MWD-survey, but random between different MWD-surveys. Let ε_4, ε_5 and ε_6 denote the accelerometer scale errors. The errors ε_4, ε_5 and ε_6 are assumed independent. From (2.2)-(2.4) we see that accelerometer errors caused by the accelerometer scale errors are:

$$\mathrm{d}G_x = G_x\varepsilon_4 = [-G\sin I \sin\tau]\varepsilon_4$$
$$\mathrm{d}G_y = G_y\varepsilon_5 = [-G\sin I \cos\tau]\varepsilon_5$$
$$\mathrm{d}G_z = G_z\varepsilon_6 = [G\cos I]\varepsilon_6$$

The value G is the gravity reference value shown in Section 2.4. Alternatively one could use $G = \sqrt{G_x^2 + G_y^2 + G_z^2}$, but then the scale error would be dependent of the bias error. For simplicity, we assume that G is the gravity reference value and that the accelerometer scale errors are independent of the accelerometer bias errors. An error dG_x in the X-accelerometer causes the errors dI and dA in I and A. If we assume that the errors are small, we can approximate dI and dA from (2.5) and (2.11):

$$\mathrm{d}I \approx \frac{\partial I}{\partial G_x}\mathrm{d}G_x = \frac{\partial I}{\partial G_x}[-G\sin I \sin\tau]\varepsilon_4 \tag{4.6}$$

$$\mathrm{d}A \approx \frac{\partial A}{\partial G_x}\mathrm{d}G_x = \frac{\partial A}{\partial G_x}[-G\sin I \sin\tau]\varepsilon_4 \tag{4.7}$$

It can be shown that the weighting functions for the accelerometer scale errors are as follows:

$$w_4 = \begin{bmatrix} 0 \\ \sin I \cos I \sin^2 \tau \\ -[\tan\theta \sin I(\cos I \sin A_m \sin\tau - \cos A_m \cos\tau) \\ + \cos I \cos\tau] \sin\tau \end{bmatrix} \tag{4.8}$$

$$w_5 = \begin{bmatrix} 0 \\ \sin I \cos I \cos^2 \tau \\ -[\tan\theta \sin I(\cos I \sin A_m \cos\tau + \cos A_m \sin\tau) \\ - \cos I \sin\tau] \cos\tau \end{bmatrix} \tag{4.9}$$

$$w_6 = \begin{bmatrix} 0 \\ -\sin I \cos I \\ \tan\theta \sin I \cos I \sin A_m \end{bmatrix} \tag{4.10}$$

4.2.3 Magnetometer bias errors

The magnetometer bias errors affect the 3 magnetometer sensors in the same way the accelerometer bias errors affect the accelerometer sensors. The magnetometer bias errors are systematic within each MWD-survey, but random between different MWD-surveys. Let ε_7, ε_8 and ε_9 denote the magnetometer bias errors. The errors ε_7, ε_8 and ε_9 are assumed to be independent. The magnetometer errors caused by the magnetometer bias errors are:

$$\begin{aligned} \mathrm{d}B_x &= \varepsilon_7 \\ \mathrm{d}B_y &= \varepsilon_8 \\ \mathrm{d}B_z &= \varepsilon_9 \end{aligned}$$

The magnetometer measurements are only used in computation of the azimuth, and errors in magnetometers do not affect the inclination, as were the case for the accelerometers. We have the following weighting functions for the magnetometer bias errors:

$$w_7 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{B\cos\theta}(\cos A_m \cos\tau - \cos I \sin A_m \sin\tau) \end{bmatrix} \tag{4.11}$$

$$w_8 = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{B\cos\theta}(\cos A_m \sin\tau + \cos I \sin A_m \cos\tau) \end{bmatrix} \tag{4.12}$$

$$w_9 = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{B\cos\theta}(\sin I \sin A_m) \end{bmatrix} \tag{4.13}$$

4.2.4 Magnetometer scale errors

A magnetometer scale error affect the magnetometer with an error proportional to the amplitude of the measurement of the sensor. The magnetometer scale errors are assumed to be systematic within each MWD-survey, but random between different MWD-surveys. Let ε_{10}, ε_{11} and ε_{12} denote the magnetometer scale errors. From (2.7)-(2.9) we see that magnetometer errors caused by the magnetometer scale errors are:

$$
\begin{aligned}
\mathrm{d}B_x &= B_x\varepsilon_{10} = B\big([\cos\theta\cos I\cos A_m - \sin\theta\sin I]\sin\tau + \cos\theta\sin A_m\cos\tau\big)\varepsilon_{10} \\
\mathrm{d}B_y &= B_y\varepsilon_{11} = B\big([\cos\theta\cos I\cos A_m - \sin\theta\sin I]\cos\tau - \cos\theta\sin A_m\sin\tau\big)\varepsilon_{11} \\
\mathrm{d}B_z &= B_z\varepsilon_{12} = B\big(\cos\theta\sin I\cos A_m + \sin\theta\cos I\big)\varepsilon_{12}
\end{aligned}
$$

The values B and θ are magnetic reference values shown in Section 2.3. We have the following weighting functions for the magnetometer scale errors:

$$
w_{10} = \begin{bmatrix} 0 \\ 0 \\ (\cos I\cos A_m\sin\tau - \tan\theta\sin I\sin\tau + \sin A_m\cos\tau) \\ \cdot(\cos A_m\cos\tau - \cos I\sin A_m\sin\tau) \end{bmatrix} \tag{4.14}
$$

$$
w_{11} = \begin{bmatrix} 0 \\ 0 \\ -(\cos I\cos A_m\cos\tau - \tan\theta\sin I\cos\tau - \sin A_m\sin\tau) \\ \cdot(\cos A_m\sin\tau + \cos I\sin A_m\cos\tau) \end{bmatrix} \tag{4.15}
$$

$$
w_{12} = \begin{bmatrix} 0 \\ 0 \\ -(\sin I\cos A_m + \tan\theta\cos I)\sin I\sin A_m \end{bmatrix} \tag{4.16}
$$

4.2.5 Depth errors

We use three independent error sources concerning the measured depth D. The depth errors are quite complex, and details concerning these errors can be found in Ekseth [9]. Let ε_{13} be the depth error concerning the reference depth. The relative distance between the reference and the wellbore is affected by many error sources, such as water tides and varying rig ballast. It is sufficient to include all these errors into ε_{13}. This error source is assumed to be randomly distributed between the measurement stations in the well. The second depth error is the scale factor, ε_{14}. It is assumed to be proportional to the measured depth. The scale errors are assumed to be systematic within each MWD-survey, and random between different MWD-surveys. The third depth error is the stretch type error, ε_{15}. The stretch type error is mainly caused by stretch in the drillstring, and thermal expansion. It is assumed to be proportional to the product of the measured depth and the vertical depth. It is considered to be systematic for all wells, i.e. global. This error dominates the other depth errors in deep wells.

The depth errors affect the measured depth as follows:

$$\begin{aligned} \mathrm{d}D &= \varepsilon_{13} \\ \mathrm{d}D &= D \cdot \varepsilon_{14} \\ \mathrm{d}D &= D \cdot D_V \cdot \varepsilon_{15} \end{aligned}$$

, where D is the measured depth and D_V is the vertical depth. The weighting functions for the depth errors are quite simple:

$$w_{13} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{4.17}$$

$$w_{14} = \begin{bmatrix} D \\ 0 \\ 0 \end{bmatrix} \tag{4.18}$$

$$w_{15} = \begin{bmatrix} D \cdot D_V \\ 0 \\ 0 \end{bmatrix} \tag{4.19}$$

4.2.6 Declination errors

The uncertainty in the magnetic declination angle is probably the most important uncertainty in the error model. As described in Section 2.3, the Earth's magnetic field can be roughly divided into three parts. These three magnetic fields vary both in position and time. It is reasonable to include the errors from the magnetic fields into our error model. The errors of the field strength B and the dip angle θ are not considered to be of significant magnitude. But the declination is considered to be the most significant error source regarding the position uncertainty in MWD magnetic directional surveying. We separate the declination errors in two parts: The constant declination error ε_{16}, and the B_H-dependent declination error ε_{17}. Both error sources are considered to have a global propagation mode. They affect the azimuth as follows:

$$\begin{aligned} \mathrm{d}A &= \varepsilon_{16} \\ \mathrm{d}A &= \frac{1}{B\cos\theta} \cdot \varepsilon_{17} \end{aligned}$$

The weighting functions are:

$$w_{16} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{4.20}$$

$$w_{17} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{B\cos\theta} \end{bmatrix} \tag{4.21}$$

4.2.7 Sag error

The sag error is caused by misalignment of the sensors in the well. The weight of the drillstring causes a sag when a well is drilled with a high inclination angle. The sag error ε_{18} affect the inclination as follows:

$$\mathrm{d}I = \sin I \varepsilon_{18}$$

The weighting function is:

$$w_{18} = \begin{bmatrix} 0 \\ \sin I \\ 0 \end{bmatrix}$$

4.2.8 Axial magnetism error in drillstring

Steel in other equipments than the MWD-tool will always cause a magnetic field around the MWD-tool. This magnetic field can be divided into an axial- and a cross-axial component. We include only the axial component in our error model. The use of non-magnetic spacing around the MWD-tool is often required in order to have control of the axial magnetic field. The axial magnetism error ε_{19} affects the azimuth as follows:

$$\mathrm{d}A = \frac{\sin I \sin A_m}{B \cos \theta} \cdot \varepsilon_{19}$$

The weighting function is:

$$w_{19} = \begin{bmatrix} 0 \\ 0 \\ \frac{\sin I \sin A_m}{B \cos \theta} \end{bmatrix}$$

4.2.9 Excluded errors

The errors that are mentioned so far are believed to be the most significant error sources in MWD magnetic directional surveying. There are however other errors that can be included in an error model as easily as the errors that are included. Random measurement errors on the MWD-tool could have been included, but they are not considered to be of significant magnitude [9]. The sag error is only one type of misalignment error. The MWD-tool can be misaligned radially to the drilling direction. The cross-axial magnetism error caused by other equipments can also be included in the model.

The non-magnetic spacing between the MWD-tool and other instruments is not always practiced. The axial magnetic field will then affect the axial measurement B_z, making B_z too inaccurate. It is possible to correct for the axial magnetism by ignoring the axial measurement B_z, and introducing the magnetic reference values from Section 2.3. This leads to a a new set of equations as in Section 2.5, which leads to a new set of weighting functions that has to be applied. Since the main focus of this thesis is not the axial magnetism error, we have not bothered including this new set of weighting functions in our model as it would complicate the model implementation. For further details about these weighting functions, see [25].

4.3 Error propagation theory

Section 4.2 introduces the weighting functions, which show how the error sources cause errors in the measured depth D, inclination angle I and azimuth angle A. In this section we show how this cause errors in the position uncertainty, and we derive a framework for computing the position uncertainty distribution if the error sources are normally distributed. It is natural to distinguish the error sources between the different propagation modes. Let $R = \{13\}$, $S = \{1, 2, .., 12, 14, 18, 19\}$ and $G = \{15, 16, 17\}$ denote the sets of random-, systematic- and global errors, respectively.

Let $w_{R,k}$ denote the weighting function of a random error ε_R at measurement station k in the well. The accumulated error contribution $[\mathrm{d}D, \mathrm{d}I, \mathrm{d}A]^T_{R,k}$ from the random error ε_R at measurement station k is:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{R,k} = \sum_{i=1}^{k} w_{R,i}\varepsilon_{R,i}$$

We assume that $\mathrm{Cov}(\varepsilon_{R,i}, \varepsilon_{R,j}) = 0$ for $i \neq j$, since the error has a random propagation mode. If we assume that $\varepsilon_{R,i} \sim N(0, \sigma_R^2)$, we can calculate the statistical distribution for the error contribution:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{R,k} \sim N_3\left(0, \sigma_R^2 \sum_{i=1}^{k} w_{R,i} w_{R,i}^T\right) \tag{4.22}$$

For simplicity, we write $C_{R,k} = \sigma_R^2 \sum_{i=1}^{k} w_{R,i} w_{R,i}^T$ as the covariance matrix at station k caused by error ε_R.

In the same way, we can analyze the error contribution from a systematic error source ε_S. Let $w_{S,l,k}$ denote the weighting function of a systematic error ε_S in MWD-survey l at measurement station k. Let n_l be the number of stations in MWD-survey l. The accumulated error contribution $[\mathrm{d}D, \mathrm{d}I, \mathrm{d}A]^T_{S,L,k}$ from the systematic error ε_S in MWD-survey L at measurement station k is:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{S,L,k} = \sum_{l=1}^{L-1} \sum_{i=1}^{n_l} w_{S,l,i}\varepsilon_{S,l} + \sum_{i\in L}^{k} w_{S,L,i}\varepsilon_{S,L}$$

We assume that $\mathrm{Cov}(\varepsilon_{S,l}, \varepsilon_{S,o}) = 0$ for $l \neq o$, since the error has a systematic propagation mode. For simplicity, we write $C_{S,l} = \sigma_S^2 \left(\sum_{i=1}^{n_l} w_{S,l,i}\right) \left(\sum_{i=1}^{n_l} w_{S,l,i}\right)^T$

and $\bar{C}_{S,L} = \sigma_S^2 \left(\sum_{i\in L}^{k} w_{S,L,i}\right)\left(\sum_{i\in L}^{k} w_{S,L,i}\right)^T$. If we assume that $\varepsilon_{S,l} \sim N(0, \sigma_S^2)$, we can calculate the statistical distribution for the error contribution:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{S,L,k} \sim N_3\left(0, \sum_{l=1}^{L-1} C_{S,l} + \bar{C}_{S,L}\right) \tag{4.23}$$

For simplicity, we write $C_{S,L,k} = C_{S,k} = \sum_{l=1}^{L-1} C_{S,l} + \bar{C}_{S,L}$.

Let ε_G be a global error source, and $w_{G,k}$ denote the weighting function of the global error at measurement station k. The accumulated error contribution $[\mathrm{d}D, \mathrm{d}I, \mathrm{d}A]_{G,k}^T$ from the global error ε_G at measurement station k is:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{G,k} = \sum_{i=1}^{k} w_{G,i}\varepsilon_G$$

If we assume that $\varepsilon_G \sim N(0, \sigma_G^2)$, the statistical distribution for the error contribution is:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{G,k} \sim N_3\left(0, \sigma_G^2 \left(\sum_{i=1}^{k} w_{G,i}\right)\left(\sum_{i=1}^{k} w_{G,i}\right)^T\right) \tag{4.24}$$

For simplicity, we write $C_{G,k} = \sigma_G^2 \left(\sum_{i=1}^{k} w_{G,i}\right)\left(\sum_{i=1}^{k} w_{G,i}\right)^T$. To find the total error $[\mathrm{d}D, \mathrm{d}I, \mathrm{d}A]_k^T$ at measurement station k we can sum all the error contributions. Let $C_k = \sum_R C_{R,k} + \sum_S C_{S,k} + \sum_G C_{G,k}$, and we get the following error distribution at measurement station k:

$$\begin{bmatrix} \mathrm{d}D \\ \mathrm{d}I \\ \mathrm{d}A \end{bmatrix}_{k} \sim N_3\left(0, C_k\right) \tag{4.25}$$

Now we have the uncertainty in measured depth D, inclination I and azimuth A. We use the same approach for computing the position uncertainty. Recall the minimum curvature formula (2.15):

$$P_k = P_{k-1} + \frac{(D_k - D_{k-1})f(\kappa_k)}{2} \begin{bmatrix} \sin I_{k-1}\cos A_{k-1} + \sin I_k \cos A_k \\ \sin I_{k-1}\sin A_{k-1} + \sin I_k \sin A_k \\ \cos I_{k-1} + \cos I_k \end{bmatrix} \tag{4.26}$$

The position is dependent on the preceding position and the two last measurements of D, I and A. This means that the error contribution from one error source in the preceding and following measurement station have direct influence on the position uncertainty. The position uncertainty is obtained by linearizing the minimum curvature formula. For small dogleg angles, we can simplify the minimum curvature formula by setting $f(\kappa) = 1$. According to Williamson [25] there is no significant loss of accuracy in using this approximation when we are interested in the position uncertainty. This approximation must however not be used when calculating the position as in Section 2.6. Let

$$\Delta P_{k+1} = \frac{D_{k+1} - D_k}{2} \begin{bmatrix} \sin I_k \cos A_k + \sin I_{k+1} \cos A_{k+1} \\ \sin I_k \sin A_k + \sin I_{k+1} \sin A_{k+1} \\ \cos I_k + \cos I_{k+1} \end{bmatrix}. \tag{4.27}$$

An error source ε_i causes a position error $\mathrm{d}P_{i,k}$ at measurement station k. From the minimum curvature formula, we understand that $\mathrm{d}P_{i,k}$ is affected by the error of ε_i both on the preceding and following measurement station. We differentiate the expression in (4.27) on the k'th and $(k+1)$'th measurement station with respect to D_k, I_k and A_k, and define:

$$\begin{aligned}
R_{j,k} &= \left[\frac{\partial \Delta P_j}{\partial D_k}, \frac{\partial \Delta P_j}{\partial I_k}, \frac{\partial \Delta P_j}{\partial A_k}\right] \quad ; \quad j = k, k+1 \\
\frac{\partial \Delta P_k}{\partial D_k} &= \frac{1}{2} \begin{bmatrix} \sin I_{k-1} \cos A_{k-1} + \sin I_k \cos A_k \\ \sin I_{k-1} \sin A_{k-1} + \sin I_k \sin A_k \\ \cos I_{k-1} + \cos I_k \end{bmatrix} \\
\frac{\partial \Delta P_{k+1}}{\partial D_k} &= \frac{1}{2} \begin{bmatrix} -\sin I_k \cos A_k - \sin I_{k+1} \cos A_{k+1} \\ -\sin I_k \sin A_k - \sin I_{k+1} \sin A_{k+1} \\ -\cos I_k - \cos I_{k+1} \end{bmatrix} \\
\frac{\partial \Delta P_j}{\partial I_k} &= \frac{D_j - D_{j-1}}{2} \begin{bmatrix} \cos I_k \cos A_k \\ \cos I_k \sin A_k \\ -\sin I_k \end{bmatrix} \quad ; \quad j = k, k+1 \\
\frac{\partial \Delta P_j}{\partial A_k} &= \frac{D_j - D_{j-1}}{2} \begin{bmatrix} -\sin I_k \sin A_k \\ \sin I_k \cos A_k \\ 0 \end{bmatrix} \quad ; \quad j = k, k+1
\end{aligned}$$

$R_{j,k}$ is a 3×3 matrix, with $\frac{\partial \Delta P_j}{\partial D_k}$, $\frac{\partial \Delta P_j}{\partial I_k}$ and $\frac{\partial \Delta P_j}{\partial A_k}$ along each column. The position change $\mathrm{d}P_{i,k}$ can be written as:

$$\mathrm{d}P_{i,k} = \left[R_{k,k} + R_{k+1,k}\right] w_{i,k} \varepsilon_i$$

The position change in the last measurement station k is given by:

$$\mathrm{d}P^{\star}_{i,k} = R_{k,k} w_{i,k} \varepsilon_i \tag{4.28}$$

Now we can use the same approach as we did when we derived the uncertainty of D, I, and A. Let $\mathrm{d}P_{R,k}$ be the total error contribution at measurement station k, caused by the random error source ε_R during the whole well. It can be computed as:

$$\mathrm{d}P_{R,k} = \sum_{i=1}^{k-1} \left[R_{i,i} + R_{i+1,i}\right] w_{R,i} \varepsilon_{R,i} + R_{k,k} w_{R,k} \varepsilon_{R,k}$$

If we assume that $\varepsilon_{R,i} \sim N(0, \sigma_R^2)$, and $\mathrm{Cov}(\varepsilon_{R,i}, \varepsilon_{R,j}) = 0$ for $i \neq j$, we have the following distribution for $\mathrm{d}P_{R,k}$:

$$\mathrm{d}P_{R,k} \sim N\left(0, \Omega_{R,k}\right)$$

$$\Omega_{R,k} = \left[\sum_{i=1}^{k-1} \left[R_{i,i} + R_{i+1,i}\right] w_{R,i} w_{R,i}^T \left[R_{i,i} + R_{i+1,i}\right]^T + R_{k,k} w_{R,k} w_{R,k}^T R_{k,k}^T\right] \sigma_R^2$$

Let $\mathrm{d}P_{S,L,k}$ be the total error contribution at measurement station k in MWD-survey L, caused by the systematic error source ε_S. It can be computed as:

$$\begin{aligned}\mathrm{d}P_{S,L,k} = & \sum_{l=1}^{L-1} \sum_{i=1}^{n_l} \left[R_{l,i,i} + R_{l,i+1,i}\right] w_{S,l,i} \varepsilon_{S,l} + \\ & + \left[\sum_{i \in L}^{k-1} \left[R_{L,i,i} + R_{L,i+1,i}\right] w_{S,L,i} + R_{L,k,k} w_{S,L,k}\right] \varepsilon_{S,L}\end{aligned}$$

The distribution of $\mathrm{d}P_{S,L,k}$ can be computed as for $\mathrm{d}P_{R,k}$, and we denote the resulting covariance matrix of $\mathrm{d}P_{S,L,k}$ for $\Omega_{S,k}$.

Let $\mathrm{d}P_{G,k}$ be the total error contribution at measurement station k, caused by the global error source ε_G. It can be computed as:

$$\mathrm{d}P_{G,k} = \left[\sum_{i=1}^{k-1} \left[R_{i,i} + R_{i+1,i}\right] w_{G,i} + R_{k,k} w_{G,k}\right] \varepsilon_G \tag{4.29}$$

The distribution of $\mathrm{d}P_{G,k}$ can be computed as for $\mathrm{d}P_{R,k}$ and $\mathrm{d}P_{S,L,k}$, and we denote the resulting covariance matrix of $\mathrm{d}P_{G,k}$ for $\Omega_{G,k}$. The total error contribution $\mathrm{d}P_k$ at measurement station k, caused by all the error sources can now be computed. The position covariance matrix at measurement station k is simply:

$$\Omega_k = \sum_R \Omega_{R,k} + \sum_S \Omega_{S,k} + \sum_G \Omega_{G,k}$$

If we assume that $P_{k-1} \sim N_3(p_{k-1}, \Omega_{k-1})$, we have from (4.26) that the expected position $p_k = p_{k-1} + \Delta p_k$. And with the covariance matrix Ω_k, derived from linearization of ΔP_k, we have approximated the position distribution in the k'th measurement station as:

$$P_k \sim N_3(p_k, \Omega_k)$$

There are some choices which have to be made when implementing the framework in Section 4.3. We have to decide whether or not there shall exist an initial uncertainty. We had a special case for $\mathrm{d}P_{i,K}^{\star}$ at the last measurement station. We have the same situation at the first measurement station, concerning the inital uncertainty. In our implementation of this model, we have set zero initial position uncertainty.

The uncertainty analysis in this section is often used in well-planning. A planned well is described by D, I and A. But some of the weighting functions need the value for the toolface angle τ to be fully evaluated. τ is unknown during well-planning since the wellbore rotate during drilling, and τ tells how the wellbore is oriented when the sensors have made the measurements. There are two ways to overcome this problem. The first alternative is to draw τ uniformly in the uncertainty analysis. The computed position uncertainties would then be random between two analyses. The uncertainty evolving the randomization of τ is not significant compared to the total position uncertainty, and can be negligible. But it is important to know about this process since the numerical results vary between analyses of the same well which apparently seem to be identical. The other possibility is to eliminate τ from the weighting functions. This method is described in [21]. We have implemented the first alternative, where we draw τ uniformly on every measurement station.

Chapter 5

The normal inverse gaussian distribution in MWD magnetic directional surveying

5.1 Introduction

This chapter introduces the use of the normal inverse gaussian distribution on the error sources. Simulations are used to analyze the position uncertainty in MWD magnetic directional surveying against the established industry standard from Chapter 4, and an approximation of the new position uncertainty is presented.

The normal inverse gaussian distribution is a flexible distribution which has 4 parameters $(\mu, \rho, \alpha, \beta)$. The four parameters are connected with the first four standardized moments through (3.13)-(3.16), which allow us to set the skewness and kurtosis almost independently of the mean and variance of the distribution. Williamson [26] suggested a "candidate distribution" which describes the tails of MWD error-distributions better than the normal distribution. This distribution $f(x) \propto \frac{1}{1+x^4}$ allows no flexibility regarding the standardized moments and parameter specification, so it is impossible to use in the way we intend do. But the work of [26] indicates that the use of a more heavy-tailed distribution than the normal distribution is needed for directional surveying.

In Chapter 4 we derived a method for computing the position uncertainty under the assumption that the error sources were normally distributed. The normal distribution is closed under linearity, and there are simple formulas for computing the mean and variance of linear combinations of normally distributed random variables. When we assumed that the error sources were normally distributed, we could use these properties of the normal distribution to compute a normally distributed position uncertainty. If we choose to use other statistical distributions for the error sources we can not use the same approach as in Chapter 4 as we lose the nice properties which we had when we used the normal distribution for all the error sources. This forces us to use other approaches than to analytically compute the position uncertainty. The weighting functions that are defined in Section 4.2 are however unchanged, and are used extensively in this chapter.

Section 5.2 explains how we can set the parameters of a NIG-distribution by interpreting the NIG-parameters. We propose two different NIG-distributions for the declination error; a symmetric NIG-distribution and a skewed NIG-distribution. The latter one is used in this thesis, while the first model is a motivation for heavy-tailed distributions. Section 5.3 shows how we can calculate an approximative density function if we assume that one of the error sources are NIG-distributed and the rest are normally distributed. The approximative density function is useful as we are able to evaluate the probability distribution. In Section 5.4 we explain how to simulate the position uncertainty given any statistical distribution on the error sources. Simulations are very important for analyzing the new position uncertainties. Section 5.5 shows all the analytical calculations we are able to do if some of the error sources are NIG-distributed. Section 5.6 shows one special well geometry which allows us to calculate analytically even further. Section 5.7 compares the approximative density function from Section 5.3 against the established normal error model from Chapter 4.

5.2 Specification of NIG-parameters for the declination error

In Section 4.2 we defined an error model which includes the most important error sources. Table 4.1 in Section 4.2 lists the different error sources, with a suitable standard deviation, i.e. $\varepsilon_i \sim N(0, \sigma_i^2)$. If we want to use a skewed- and heavy-tailed distribution, such as the NIG distribution, on one or more of the error sources $\varepsilon_i \sim NIG(\mu_i, \rho_i, \alpha_i, \beta_i)$, we must find a way to set the parameters $(\mu_i, \rho_i, \alpha_i, \beta_i)$ of the distribution. The ideal approach would be to estimate the parameters according to Section 3.4.2, but it is hard to obtain datasets of the error sources. Estimates of the bias- and scale-errors from multi-station estimation [17] is a possible way to obtain data of the errors ε_1-ε_{12} in Table 4.1. Nyrnes [17] shows some histograms of these errors, which could indicate that they are non-normally distributed, although these histograms are based on relatively few datapoints. The declination error ε_{16} is possible to analyze from data of the Earth's magnetic field, which is very easy to obtain datasets from. One problem is that the declination error is highly dependent both on the time and the geographic position, [22] and [23].

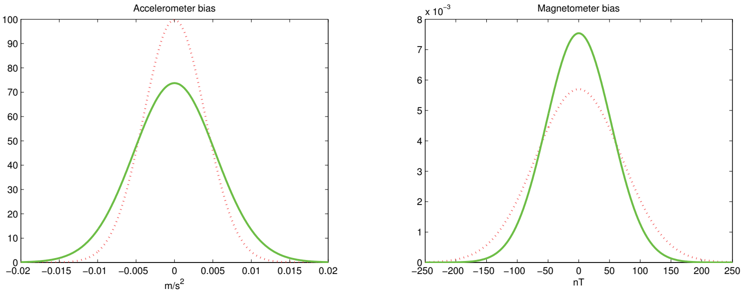


Figure 5.1: Choice of NIG-distribution for two of the error sources. The dashed red lines are the normal distributions proposed in the error model in Section 4.2. Left: $\varepsilon_1 \sim NIG(0, 0.03, 1000, 0.07)$. Right: $\varepsilon_7 \sim NIG(-196, 2800000, 1000, 0.07)$.

For the error sources which no data can be obtained from, the parameters of the NIG-distribution must be set in a way which reflects the desired properties of the NIG-distribution. The shape of the NIG-distribution can not be immediately interpreted from given parameters, as for the normal distribution. The μ and σ in the normal distribution tells us the location and scale of the distribution. If the scale of a NIG-distribution is large, it would imply that we set the scale parameter ρ large. If we want the distribution to be centered around 0 ($E(\varepsilon) = 0$), we see from (3.13) that we must set $\mu = -\frac{\rho\beta}{\gamma}$.

One way to set μ and ρ is to solve $E(\varepsilon) = 0$ and $\mathrm{Var}(\varepsilon) = \sigma^2$ from (3.13)-(3.14), where σ is the suggested standard deviation of ε from Table 4.1. The skewness is determined by the sign and size of β. The kurtosis is determined by the size of α and β together. Of (3.16) we see that if α is small relative to β, or if α and β are close such that γ is small, the kurtosis becomes large, and the distribution becomes very peaked and heavy-tailed. If $\alpha \gg \beta$, $\beta \neq 0$, and μ balanced compared to ρ, we get distributions which are skewed and more heavy-tailed than the normal distribution. Let $\Lambda_i = [\mu_i, \rho_i, \alpha_i, \beta_i]$ be the set of parameters for the NIG-distribution for $\varepsilon_i \sim NIG(\mu_i, \rho_i, \alpha_i, \beta_i)$. Figure 5.1 shows examples of two different NIG-distributions with the same choice of α and β. But since the magnetometer bias error has a proposed standard deviation of 70 nT, and the accelerometer bias error has a proposed standard deviation of 0.0039 m/s^2, the chosen μ and ρ are different.

We now introduce two different NIG-distributions for the declination error ε_{16}. The declination error is probably the largest error source in MWD magnetic directional surveying, which [22] and [23] show is likely to follow a skewed- and heavy-tailed distribution.

5.2.1 Symmetric NIG-distributed declination error

A report of Torkildsen, Sveen and Bang [23] shows a study of the variation of the declination angle in the geomagnetic field. The main result of [23] is that the statistical distribution of the declination error deviates strongly from the normal distribution. The empirical 99.5-percentiles deviate with a factor 2, and the empirical 99.9-percentiles deviate with a factor 3. The report does not focus on the skewness of the statistical distribution. If we want to model the declination error with a heavy-tailed symmetric distribution, we can use a symmetric NIG-distribution. It is reasonable for us to set $\mu = \beta = 0$, which makes $E(\varepsilon_{16}) = 0$ and $\mathrm{Skew}(\varepsilon_{16}) = 0$. We choose ρ and α from $\mathrm{Var}(\varepsilon_{16}) = \sigma_{16}^2$ and the main result from [23] regarding the percentiles. σ_{16} is the suggested standard deviation in Table 4.1. If we set $\rho = 0.00275$ and $\alpha = 70$, we have a NIG-distribution which fulfils the main result of [23].

5.2.2 Skewed NIG-distributed declination error

In the project thesis [10], several datasets from geomagnetic observatories [18] were analyzed. Although these datasets are both time- and location-dependent it is possible for us to center each dataset and remove any linear trend in the dataset to analyze the declination variability. We have chosen the annual declination series of Tromsø in 2001 as our dataset. It is important to make notice that the model we present is only based on data from one location in one year. We process the dataset by centering it, removing the linear trend, and scaling it such that $\mathrm{Var}(\varepsilon_{16}) = \sigma_{16}^2$. It is of interest for our thesis to analyze a non-central error source $E(\varepsilon_{16}) \neq 0$. Of the original time series with a resolution of 1 minute, we calculate the following moment estimators of a random sample of size $\frac{1}{100}$ of the processed time series:

$$\begin{aligned}
\hat{\mu} &= -5.43 \cdot 10^{-4} \\
\hat{\rho} &= 0.0021 \\
\hat{\alpha} &= 58.33 \\
\hat{\beta} &= 14.6
\end{aligned}$$

Of the reasons explained in Section 3.4.2 we have chosen to use the moment estimators as our parameter values. The moment estimators of the processed dataset will be the parameters of the skewed NIG-distribution which we use in the NIG error model in this thesis. Table 5.1 shows a summary of the choice of parameters for the NIG-distributions. Figure 5.2 displays the histogram of the processed dataset Tromsø 2001, the two NIG-distributions and the corresponding normal distribution. Figure 5.2 only show the centre of the dataset, so it is possible to see the difference between the models. We see that the NIG-distributions are more peaked than the normal distribution. But the NIG-distributions also have heavier tails than the normal distribution.

	μ	ρ	α	β	γ
Symmetric NIG	0	0.00275	70	0	70
Skewed NIG	$-5.43 \cdot 10^{-5}$	0.0021	58.33	14.6	56.47

Table 5.1: A summary of the choice of NIG-parameters for the two NIG-distributions of the declination error. The symmetric NIG-distribution is partly obtained from [22], while the skewed NIG-distribution is obtained from real processed data [10].

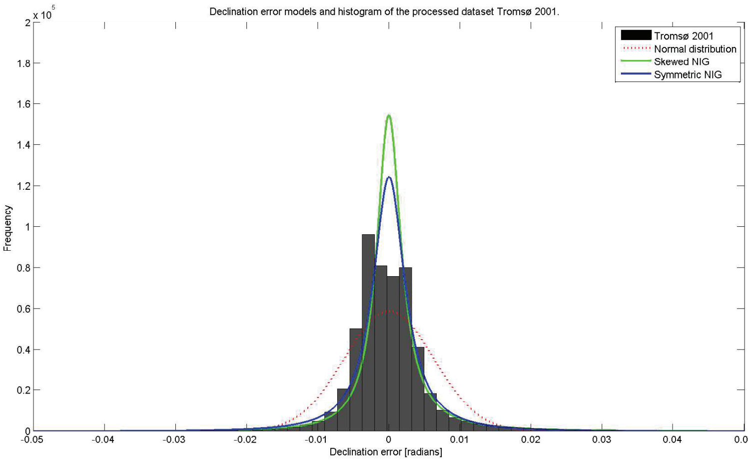


Figure 5.2: The two proposed NIG-distributions for the declination error. The blue line is the symmetric NIG-distribution, and the green line is the skewed NIG-distribution. The histogram is the centre of the processed dataset Tromsø 2001, as the range of the processed dataset is 0.26 radians. The dashed red line is the normal distribution. The NIG-parameters of the models are shown in Table 5.1.

5.3 Approximative density function

We now consider the declination error ε_{16} to be the only error that is NIG-distributed. The declination error is considered to be the most significant error source in MWD magnetic directional surveying, and it is shown in [23] that it has heavy-tailed behaviour. The other error sources (Table 4.1) are normally distributed. We can partition the position uncertainty $\mathrm{d}P_k$ in two independent parts as we assume that the error sources are independent of eachother:

$$\mathrm{d}P_k = \mathrm{d}P_{N,-16,k} + \mathrm{d}P_{NIG,16,k}$$

$\mathrm{d}P_{N,-16,k}$ is the position uncertainty at measurement station k based on all error sources except the declination error and $\mathrm{d}P_{NIG,16,k}$ is the position uncertainty at measurement station k based only on the declination error, which is NIG-distributed. We recall the following property of the NIG-distribution:

$$\begin{aligned} \varepsilon_{16} &\sim NIG(\mu, \rho, \alpha, \beta) \\ \{\varepsilon_{16} | Z = z\} &\sim N(\mu + \beta z, z) \\ Z &\sim IG(\rho, \gamma) \end{aligned}$$

We condition on the auxiliary variable Z in the position uncertainty:

$$\{\mathrm{d}P_k | Z = z\} = \mathrm{d}P_{N,-16,k} + \{\mathrm{d}P_{NIG,16,k} | Z = z\}$$

The uncertainty from the NIG-distributed error source is normally distributed when conditioned on Z. From (4.29) the following can be shown:

$$\begin{aligned} \mathrm{d}P_{NIG,16,k} &= \left(\sum_{i=1}^{k-1} \left[\frac{\partial \Delta P_i}{\partial A_i} + \frac{\partial \Delta P_{i+1}}{\partial A_i} \right] + \frac{\partial \Delta P_k}{\partial A_k} \right) \varepsilon_{16} \\ &= v_k \varepsilon_{16} \end{aligned}$$

For simpler expressions we denote the sum of vectors for v_k. The position uncertainty distribution, conditioned on the auxiliary variable Z is then:

$$\begin{aligned} \{\mathrm{d}P_k | Z = z\} &= \mathrm{d}P_{N,-16,k} + \{\mathrm{d}P_{NIG,16,k} | Z = z\} \\ &\sim N_3 \left(v_k(\mu + \beta z), \Omega_{N,-16,k} + z v_k v_k^T \right) \end{aligned}$$

$\Omega_{N,-16,k}$ is the covariance matrix calculated from the method in Section 4.3 with the declination error excluded. The calculated position at measurement station k is p_k. The statistical distribution on measurement station k is:

$$\{P_k | Z = z\} \sim N_3 \left(p_k + v_k(\mu + \beta z), \Omega_{N,-16,k} + z v_k v_k^T \right)$$

From now on, we drop the k-notation. Until now, we have conditioned on the auxiliary variable Z. We want to find the statistical distribution of P. It is defined as

$$f_P(p) = \int f_{P|Z}(p|z) f_Z(z) \mathrm{d}z$$

where $f_Z(z) \sim IG(\rho, \gamma)$. This integral is not possible to solve analytically, so we use the following numerical approximation of this integral:

$$f_P(p) \approx \sum_{z_i} f_{P|Z}(p|z_i) P(z_i) \tag{5.1}$$

$P(z_i)$ is the discretization of $f_Z(z)$, and it is important to normalize $P(z_i)$ so that $\sum_{z_i} P(z_i) = 1$. If we discretize $f_Z(z)$ with a large enough grid, the approximation converges against the true distribution. From the discretization, we are able to numerically evaluate the statistical distribution in any position. This method can in general be done for all error sources being NIG-distributed. But for each error source being NIG-distributed, we must condition on an auxiliary variable. This result in as many sums as there are NIG-distributed error sources in (5.1), which require a lot of computer time for one evaluation of the approximative density function.

5.4 Simulating the position uncertainty with the NIG error model

The simplest way to see the effect of one error source being NIG-distributed is to calculate the normal error model in Section 4.2 with the NIG-distributed error source excluded. We use the following notation:

$$P_k = P_{N,-i,k} + P_{NIG,i,k}$$

P_k is the position uncertainty at measurement station k, $P_{N,-i,k}$ is the normal distributed position uncertainty at measurement station k based on all error sources except error source nr. i. It is easily simulated once the covariance matrix is calculated. $P_{NIG,i,k}$ is the position uncertainty at measurement station k based only on error source nr. i, which is NIG-distributed. $P_{NIG,i,k}$ is not a known distribution, which implies that we must use simulations in order to study the distribution of P_k. $P_{NIG,i,k}$ is easily simulated from (4.29) if it is a global error, and we are free to choose any statistical distribution on the error sources. It can easily be done for several of the error sources being NIG-distributed as well.

We visualize the simulations in one plot containing several subplots. Figure 5.3 is an example of a simulation result. The three subplots in the right column in Figure 5.3 shows the wellpath projected onto the NE-, NV- and EV-plane. The well which is analyzed in this figure is a synthetic well. A blue star indicates the station of interest, which is the point in the well where the simulation is performed. The upper subplot in the second column from the right visualizes the NIG-distribution which is used for the error source. In this case, we have used the skewed NIG-distribution on the declination error, as in Section 5.2.2. The blue distribution shows the corresponding normal distribution from Chapter 4. Under this subplot, the parameters of the NIG-distribution is displayed, and the statistical moments are calculated. The simulations are shown in the lower triangular subplots, projected onto the NE-, NV- and EV-plane and the N-, E- and V-axis. The reader must not be confused by the fact that the N- and E- axes have switched place in the bivariate subplots compared to the subplots describing the wellpaths. The axes are set in this way to be consistent with the histogram-axes.

In the 2D-projections we have plotted a 99.9% confidence ellipse, which is computed from the marginals of the normal error model. We count the amount of realizations in the simulation that is inside different confidence ellipsoids. This can help us compare the simulations against the normal error model from Chapter 4. The 2D-projections are hard to visually compare against a multivariate normal distribution, since we use many realizations. But non-normality is easily discovered in the histograms, since we know that marginals of a multivariate normal distribution are normally distributed. In Figure 5.3 we see that the position uncertainty is highly non-normally distributed, since there are heavy tails both on the N- and E-axis.

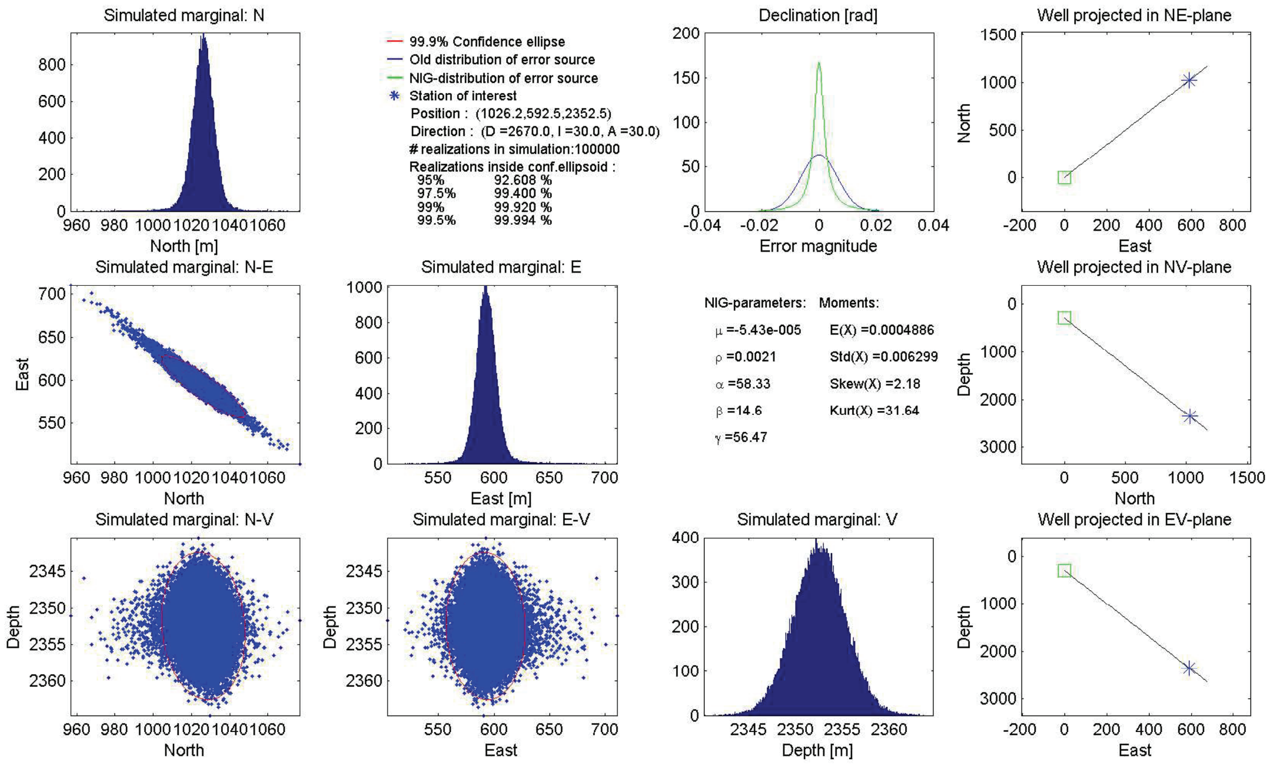


Figure 5.3: An example of a simulation plot where the declination error is NIG-distributed and the other errors are normally distributed.

5.5 Some analytical calculations

The NIG-distribution is closed under linearity only for special linear transformations. In this section we show how far it is possible to use the NIG-distribution in an analytical approach, assuming that the NIG-distribution is suitable for all error sources. Recall the weighting functions defined in Section 4.2. ($w = [w_D, w_I, w_A]^T$) For an error source ε_R of random propagation mode, we look at the error in measured depth, $\mathrm{d}D_{R,k}$, that is caused by ε_R on measurement station k:

$$\mathrm{d}D_{R,k} = \sum_{i=1}^{k} w_{D,i}\varepsilon_{R,i} \tag{5.2}$$

If we assume that $\varepsilon_{R,i} \sim NIG(\mu, \rho, \alpha, \beta)$, we know that

$$w_{D,i}\varepsilon_{R,i} \sim NIG\left(w_{D,i}\mu, |w_{D,i}|\rho, \frac{\alpha}{|w_{D,i}|}, \frac{\beta}{w_{D,i}}\right)$$

from the properties of the NIG-distribution. Since $\mathrm{Cov}(\varepsilon_{R,i}, \varepsilon_{R,j}) = 0$ for $i \neq j$, all the terms in the sum in (5.2) are independent. But the distribution of $\mathrm{d}D_{R,k}$ is in general unknown, since the NIG-distribution only is closed under linearity for a sum of independent random variables with the same skewness parameters. But there is only one value which $w_{D,i}$ can have. In Section 4.2 we defined only one error source with random propagation mode, the reference depth ε_{13}. And since $w_{13} = [1, 0, 0]^T$, we see that $w_{D,i} = 1$ for every measurement station i. This makes us able to compute the statistical distribution of the measured depth error (5.2) that is caused by the NIG-distributed random error ε_{13}:

$$\mathrm{d}D_{13,k} \sim NIG(k\mu, k\rho, \alpha, \beta) \tag{5.3}$$

For error sources with systematic propagation mode it is harder to obtain any analytical result. For an error source ε_S of systematic propagation mode, we look at the error in inclination, $\mathrm{d}I_{S,L,k}$, that is caused by ε_S on measurement station k in survey L:

$$\mathrm{d}I_{S,L,k} = \sum_{l=1}^{L-1}\sum_{i=1}^{n_l} w_{I,l,i}\varepsilon_{S,l} + \sum_{i\in L}^{k} w_{I,L,i}\varepsilon_{S,L} \tag{5.4}$$

The weighting functions for the systematic error sources are more complicated than the one for the random depth error ε_{13}. We are not able to find any element in the weighting functions for the systematic error sources which can be equal to 1 in general. For the sag error, ε_{18}, we have $w_I = \sin I$. For horizontal wells, we have that $w_I = 1$ for this error. But of $\mathrm{d}I_{S,L,k}$ in (5.4) we see that the systematic error from each survey is weighted with the number of elements n_l in each survey. We are not able to analytically compute the statistical distribution of $\mathrm{d}I_{S,L,k}$ or $\mathrm{d}A_{S,L,k}$ from any systematic error source.

For an error source ε_G of global propagation mode, we look at the error in azimuth, $\mathrm{d}A_{G,k}$, that is caused by ε_G on measurement station k:

$$\mathrm{d}A_{G,k} = \sum_{i=1}^{k} w_{A,i}\varepsilon_G \tag{5.5}$$

If $\varepsilon_G \sim NIG(\mu, \rho, \alpha, \beta)$, it is possible to analytically compute the statistical distribution of $\mathrm{d}A_{G,k}$. This computation is possible for every element in all the weighting functions for the error sources that have global propagation mode. We get the following distributions:

$$\mathrm{d}D_{15,k} \sim NIG\left(\mu \sum_{i=1}^{k} D_i D_{V,i}, \rho \sum_{i=1}^{k} D_i D_{V,i}, \frac{\alpha}{\sum_{i=1}^{k} D_i D_{V,i}}, \frac{\beta}{\sum_{i=1}^{k} D_i D_{V,i}}\right) \tag{5.6}$$

$$\mathrm{d}A_{16,k} \sim NIG\left(k\mu, k\rho, \frac{\alpha}{k}, \frac{\beta}{k}\right) \tag{5.7}$$

$$\mathrm{d}A_{17,k} \sim NIG\left(\frac{\mu k}{B\cos\theta}, \frac{\rho k}{B\cos\theta}, \frac{\alpha B\cos\theta}{k}, \frac{\beta B\cos\theta}{k}\right) \tag{5.8}$$

The distributions from (5.3), (5.6)-(5.8) are the only ones that we are able to compute analytically. We can not sum together these error contributions with other error contributions, and still know the statistical distribution of the sum. But these distributions can be useful for comparing choices of NIG-distributions against the normal distribution to see how one of these error sources contribute to the error in measured depth or azimuth.

5.6 Approximations for special well geometries

In Section 5.5 we showed some special cases where we were able to compute the distribution of the error in measured depth and azimuth from one of the error sources being NIG-distributed. For certain well geometries, it is possible to use some approximation to obtain marginal distributions of the position that are NIG-distributed. We use $\mathrm{d}A_{16}$ from (5.7) as an example. We look at the azimuth error at measurement station k and $k-1$ from the declination error ε_{16}:

$$\mathrm{d}A_{16,k-1} \sim NIG\left((k-1)\mu, (k-1)\rho, \frac{\alpha}{k-1}, \frac{\beta}{k-1}\right)$$
$$\mathrm{d}A_{16,k} \sim NIG\left(k\mu, k\rho, \frac{\alpha}{k}, \frac{\beta}{k}\right)$$

We look at the azimuth error in the deeper sections of a well, and assume that k is large so that $\frac{\alpha}{k-1} \approx \frac{\alpha}{k}$ and $\frac{\beta}{k-1} \approx \frac{\beta}{k}$. Then $\mathrm{d}A_{16,k-1}$ and $\mathrm{d}A_{16,k}$ have approximately the same statistical distribution. From linearizations of the minimum curvature formula (2.15), we have the following formulas for computing the error in the north- and east-direction at measurement station k, $\mathrm{d}N_k$ and $\mathrm{d}E_k$ respectively:

$$\mathrm{d}N_k = -\frac{D_k - D_{k-1}}{2}\left(\sin I_{k-1}\sin A_{k-1}\mathrm{d}A_{16,k-1} + \sin I_k \sin A_k \mathrm{d}A_{16,k}\right)$$
$$\mathrm{d}E_k = \frac{D_k - D_{k-1}}{2}\left(\sin I_{k-1}\cos A_{k-1}\mathrm{d}A_{16,k-1} + \sin I_k \cos A_k \mathrm{d}A_{16,k}\right)$$

The error contribution in the vertical-direction is always zero for the declination error source. The distance between two measurement stations is usually approximately 30 meters. If we assume that $D_k - D_{k-1} = 30$, and the well is straight ($I_{k-1} = I_k = I$ and $A_{k-1} = A_k = A$), and that the well geometry satisfies $\sin I \sin A = \sin I \cos A = \frac{1}{15}$, we see that $\mathrm{d}N_k = -(\mathrm{d}A_{16,k-1} + \mathrm{d}A_{16,k})$ and $\mathrm{d}E_k = \mathrm{d}A_{16,k-1} + \mathrm{d}A_{16,k}$. If all these conditions hold, we have the following approximate distributions for the marginal position uncertainties:

$$\mathrm{d}N_k \sim NIG\left(-(2k-1)\mu, (2k-1)\rho, \frac{\alpha}{k}, -\frac{\beta}{k}\right) \tag{5.9}$$

$$\mathrm{d}E_k \sim NIG\left((2k-1)\mu, (2k-1)\rho, \frac{\alpha}{k}, \frac{\beta}{k}\right) \tag{5.10}$$

Judging by all the conditions that must hold for this approximation to be valid, it is fair to believe that this approximation can only be done on synthetic wells. Of $\sin I \sin A = \sin I \cos A$, we understand that $A = 45$. For wells in this direction, an azimuth error causes an equal amount of error in the north- and east-direction, independent of the statistical distribution of the error source. The condition $\sin I \cos A = \frac{1}{15}$ implies that the theoretical inclination angle must be $I = 5.4099$. From simulations we have concluded that this approximation is only valid for $A = 45$, $I = 5.4099 \pm 1$, and $k > 100$.

5.7 Comparing the discretized approximative density function with the normal error model

The normal error model ensures that the expected position always is the position calculated from the minimum curvature formula (2.15). The skewed NIG-distribution from Section 5.2.2 causes a deviation between the calculated position from the minimum curvature formula and the expected position of the NIG error model. From the parameters of the skewed NIG-distribution in Table 5.1 we can compute the expected value $E(\varepsilon_{16}) = 4.88 \cdot 10^{-4}$ radians. Recall (4.29) from Section 4.3, for the declination error:

$$\begin{aligned}\mathrm{d}P_{16,k} &= \left[\sum_{i=1}^{k-1} \left[R_{i,i} + R_{i+1,i}\right] w_{16,i} + R_{k,k} w_{16,k}\right] \varepsilon_{16} \\ &= \left[\sum_{i=1}^{k-1} \left[\frac{\partial \Delta P_i}{\partial A_i} + \frac{\partial \Delta P_{i+1}}{\partial A_i}\right] + \frac{\partial \Delta P_k}{\partial A_k}\right] \varepsilon_{16}\end{aligned}$$

The error source ε_{16} with global propagation mode results in a total error contribution $\mathrm{d}P_{16,k}$ at measurement station k. By calculating the expected value of $\mathrm{d}P_{16,k}$, we can calculate the expected deviation of the position between the NIG error model and the normal error model. This deviation is highly dependent on the well geometry.

We want to visualize the position uncertainty from the NIG error model. This can either be done with simulations, or we can use the approximative density function (5.1). We look at the 2D-marginals of the approximative density function (5.1) and compare it against the 2D-marginals of the normal error model. We evaluate (5.1) in every point in a 3D-grid, and compute approximative 2D-marginal distributions by summing over one of the dimensions. The 2D-marginals can be displayed with contours, which can be compared against the corresponding 2D-marginals from the normal error model. Figure 5.4 shows the 2D-marginals (unscaled) of the approximative density function of a case. The contours of the 2D-marginals are highly non-normally distributed. The NE-marginal has very heavy tails along the N-coordinate. The contours of the NV- and EV-marginals are onion-shaped, with the V-marginal being normally distributed (since unchanged from Section 4.2). This case is a well that has moved relatively along the E-coordinate from the starting position, which is why the uncertainty is largest along the N-coordinate.

Figure 5.5 shows the NE-marginal of the corresponding multivariate normal distribution of the same case as Figure 5.4. By comparing Figure 5.5 with the upper plot in Figure 5.4, we see that the approximative density function has a higher mode than the multivariate normal distribution. But the approximative density function decays faster than the normal distribution around the mode. The tails of the approximative density function is much wider than the corresponding normal tails.

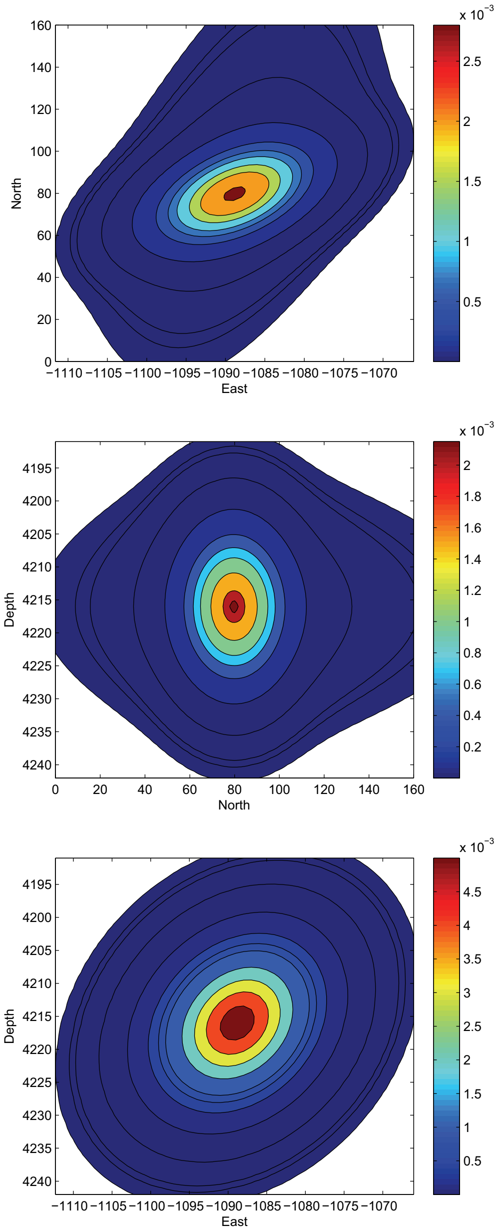


Figure 5.4: Contour plots of the approximative marginal distributions of the NIG error model. The approximative density function is evaluated on a 3D-grid and the 2D-marginals are computed by summing over the last dimension. The contour lines are manually chosen. We see that the contours of the approximative density function is non-normal, especially in the tails.

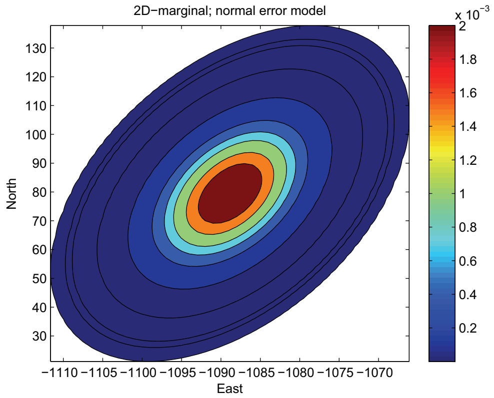


Figure 5.5: The NE-marginal of the corresponding multivariate normal distribution from the normal error model of anti-collision case II. The figure is comparable with the upper plot in Figure 5.4, and the contour lines have the same values as in Figure 5.4.

Chapter 6

Anti-collision calculations

6.1 Introduction

Well-collisions are a major risk to casualties, environmental- and economical costs, such as damage of equipment and production down-time. In order to avoid well-collisions there are different types of collision avoidance techniques that can be applied during the well-planning process. We use a quantitative anti-collision method involving a separation factor ω. Other anti-collision methods are described in [15]. The separation factor is a calculated test statistics in a statistical hypothesis test where the significance level is the tolerance probability of collision. We use a significance level $\alpha = \frac{1}{500} = 0.002$ in this thesis. In the anti-collision setting we denote the planned well as the reference well. All the existing wells around the reference well are offset wells. In this thesis we look at cases where we have only one offset well. The analyses are without loss of generality with only one offset well, as the separation factor is calculated between a point in reference well and the closest point in any of the offset wells, which is called the candidate point. The candidate point is found by a chosen scanning method. Since the separation factor is calculated between two points, it is needed to calculate a separation factor for many points along the reference well. If any of the calculated separation factors result in probabilities of collision larger than the tolerance probability, the reference well has to be replanned in a way such that all the calculated separation factors result in probabilities of collision smaller than the tolerance probability.

Section 6.2 presents the most common scanning methods that can be used to find a candidate point in the offset wells. Section 6.3 gives a brief comparison of the different scanning methods. It is important that the same scanning method is used on all anti-collision calculations in order to have consistent results. We have chosen the 3D Closest approach as our default scanning method. Section 6.4 presents the anti-collision theory that the separation factor is based upon. We present two different hypothesis tests in Section 6.4.1 and 6.4.2. Test I in Section

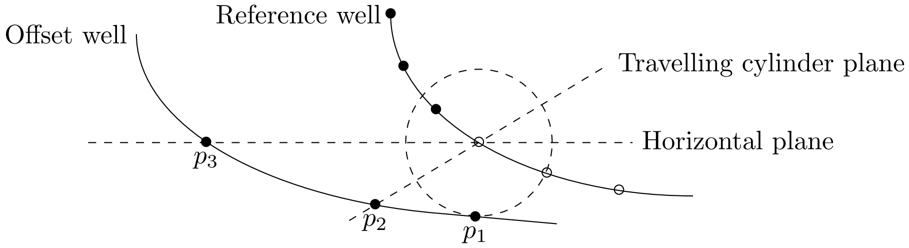


Figure 6.1: Schematic illustration of the different scanning methods which are commonly used for classifying the closest point in an offset well. The filled points in the reference well are points where anti-collision calculations have been performed. p_1 is the candidate point of the 3D Closest approach method, p_2 is the candidate point of the travelling cylinder method, and p_3 is the candidate point of the horizontal plane method.

6.4.1 is the established industry standard today, while test II in Section 6.4.2 is an empirical hypothesis test which we have to use when we use simulations of the NIG error model. In Section 6.5 we highlight the term probability of collision, and discuss how this probability is connected with the hypothesis test.

6.2 Scan methods

Anti-collision calculations are done in several points in the reference well, which we denote scanning points. In order to do anti-collision calculations we are dependent on knowing how much the offset wells are separated from the scanning point in the reference well. The anti-collision calculations are always done with respect to the "closest" point in an offset well. The distance between the scanning point in the reference well and the closest point in the offset well is called the centre to centre distance. There are several scanning methods for determining the closest point in an offset well. The different scanning methods result in different results of the anti-collision calculations, which implies that anti-collision calculations done with different scanning methods are incomparable. We present four different scanning methods which, in general, lead to different candidate points in the offset well. Details of these methods are found in [14]. The 3D Closest approach scanning method in Section 6.2.1 is our default scanning method. This is the most intuitive scanning method, and the only method that is guaranteed to find a candidate point in an offset well. Figure 6.1 illustrates the difference between the scanning methods.

6.2.1 3D Closest approach

The 3D closest approach is the most intuitive method of determining the closest point on an offset well. This method calculates the minimal Euclidean distance between the scanning point and any point in the offset wells. This method can be thought of as an expanding sphere centered in the scanning point. The sphere stops expanding when a point in an offset well is initially touched. This point is the closest point in the offset well. The 3D Closest approach method is the only scanning method that guarantees that a candidate point in the offset well will be found. This is the most commonly used scanning method in the oil and gas industry, and we also do our analyses with this scanning method.

6.2.2 Travelling cylinder

The travelling cylinder method is dependent on path of the reference well. This method uses the plane orthogonal to the wellpath, the travelling cylinder plane, to find the closest point in the offset well. We can calculate the minimal distance in the travelling cylinder plane to the intersection of some offset well. This intersection defines the closest point. The distance between the scanning point and the closest point can be thought of as a maximum scanning radius. Since there are many scanning points along the reference wellpath, we can imagine a cylinder which moves along the wellpath with varying radius, which is the reason why the method is called the travelling cylinder method. Note that this method does not guarantee that we find a candidate point in the offset well. This occurs if the offset well never crosses the travelling cylinder plane. If we include a statistical model of the position uncertainty into the scanning methods, we are able to do a modification of this method. The uncertainty in the scanning point is usually represented by a covariance matrix. The two largest eigenvectors of the covariance matrix (which are usually almost orthogonal to the wellbore direction) can be used to define a modified travelling cylinder plane.

6.2.3 Modified travelling cylinder

This method is also called ”High side + Azimuth” in some literature, but we have not chosen to use this name, as the method is a modification of the travelling cylinder method. The modified travelling cylinder uses the plane orthogonal to the wellpath in the same way as the travelling cylinder method. The only difference is that this method defines the toolface orientation as the high-side angle plus the direction of the wellpath. This method is prefered instead of the travelling cylinder method for wells with low inclination angles where the azimuth angle changes rapidly. For more details concerning this method, see [14]. We can modify this method also in the same way as the travelling cylinder method by defining the plane from the two largest eigenvectors of the covariance matrix in the scanning point.

Scan method		Comments
3D Closest approach	+	The true closest point is found, unaffected by the well geometries.
	+	Guaranteed to find a candidate point.
	−	Can find a point that is behind the reference well.
Travelling cylinder	+	Finds the closest point ahead of the reference well.
	−	The same candidate point can be found from several scanning points.
	−	Misinterpretation when the reference well crosses vertical [14].
Modified travelling cylinder	+	Avoids the misinterpretation of the travelling cylinder diagram.
	−	Hard to understand the modification.
Horizontal plane	+	Easy to understand.
	−	Should only be used for (almost) vertical wells.

Table 6.1: Summary of the advantages (+) and disadvantages (−) of the different scanning methods.

6.2.4 Horizontal plane

The horizontal plane method calculates the horizontal distance from the path of the reference well to the offset wells. The shortest distance defines the closest point. This method is equivalent to the travelling cylinder method, except that the cylinder expands horizontally instead of orthogonal to the wellbore direction. This method is only recommended to use for wells with low inclination angles. Neither this method is guaranteed to find a candidate point in the offset well. If two wells are drilled horizontally paralell to eachother, this method is not able to compute a closest point. The method is neither applicable for reference wells that are drilled deeper than the offset wells.

6.3 Comparison of the scanning methods

We have summarized some of the advantages and disadvantages of the different scanning methods in Table 6.1. Most of this table is found in [14]. It is important that one scanning method is chosen for all analyses, so the anti-collision results can be comparable. We have chosen the 3D Closest approach as our default scanning method throughout the thesis.

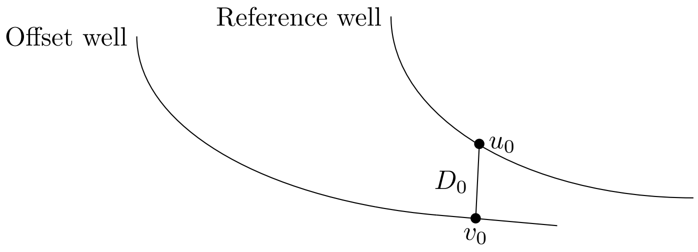


Figure 6.2: Schematic illustration of the calculated coordinate u_0 in the reference well, the calculated coordinate v_0 in the offset well, and the calculated minimum centre to centre distance D_0.

6.4 Anti-collision hypothesis testing

When we have found the closest point in an offset well and we have chosen a statistical model, we must check if there is a significant risk of collision between the offset well and the planned reference well. This is initially done with a hypothesis test, which allows us to control the type I error; the probability of accepting a wellplan given that we have a well-collision. We set the significance level $\alpha = \frac{1}{500} = 0.002$. At first, we look into the theory which is currently used today by the oil and gas industry.

We assume that the reference well and the offset well are uncorrelated, as it is currently practised in the industry standard. This is in general not the case, as global error sources influence both wells. The correlation between the wells are very dependent on the well geometries, as we can see of (4.29). The wells are often similar in some sections. Examples are when both wells are drilled from the same seabed template, or when the reference well is a sidestep of the offset well. Let u_0 be the calculated coordinate of the scanning point in the reference well, and v_0 be the calculated coordinate of the closest point in an offset well. Likewise, let u and v denote the true positions of the points. Let r be the difference vector between u and v, i.e. $r = u - v$. Likewise, let $r_0 = u_0 - v_0$ denote the calculated difference vector. Let us first assume that the positions are multivariate normally distributed, as in Section 4.3:

$$u \sim N_3(u_0, \Sigma_u)$$
$$v \sim N_3(v_0, \Sigma_v)$$

Σ_u and Σ_v are the covariance matrices of u and v, respectively. From the properties of the multivariate normal distribution, we can obtain a statistical distribution of r:

$$r \sim N_3(r_0, \Sigma_r) \quad ,$$

where $\Sigma_r = \Sigma_u + \Sigma_v$, since we assumed that the wells are uncorrelated. The distance D between u and v is defined as the Euclidean distance between u and v, i.e. $D = \sqrt{r^T r}$. The calculated distance is then $D_0 = \sqrt{r_0^T r_0}$. Figure 6.2

shows a schematic illustration of u_0, v_0 and D_0. Since we know the distribution of r, we are able to approximate the variance of D. Assume that small changes $\mathrm{d}u$ and $\mathrm{d}v$ occur in u and v. Since the changes are small, we can use a linear approximation to see how these changes result in a change $\mathrm{d}D$ in D:

$$\begin{aligned}\mathrm{d}D &\approx \frac{\partial D}{\partial u}\mathrm{d}u + \frac{\partial D}{\partial v}\mathrm{d}v \\ &= \frac{r^T}{\sqrt{r^T r}}\mathrm{d}u - \frac{r^T}{\sqrt{r^T r}}\mathrm{d}v \\ &= \frac{r^T}{\sqrt{r^T r}}[\mathrm{d}u - \mathrm{d}v]\end{aligned}$$

We compute an approximation of the variance σ_D^2 of D:

$$\begin{aligned}\sigma_D^2 &\approx \frac{r^T}{\sqrt{r^T r}}\mathrm{Cov}(\mathrm{d}u - \mathrm{d}v)\left(\frac{r^T}{\sqrt{r^T r}}\right)^T && (6.1)\\ &= \frac{1}{r^T r} r^T\left(\Sigma_u + \Sigma_v\right) r && (6.2)\end{aligned}$$

We can perform statistical hypothesis testing on the distance D. For every point in the reference well, we have a maximum tolerance probability, which is the significance level α, of accepting the wellplan given that the wellplan results in a well-collision. We set the null hypothesis to be a well-collision, versus a non-collision as the alternative hypothesis. The wellplan can not be realized if we can not reject all null hypotheses for all the points in the reference well. We now present two different hypothesis tests which can be used in anti-collision calculations. Test I in Section 6.4.1 is the current practice in the oil and gas industry. Test II in Section 6.4.2 is a generalization of test I which uses empirical distributions from simulations. An alternative hypothesis test on the difference vector r is shown in Appendix D, but this test is not applicable when the wellbore diameters are included in the model.

6.4.1 Hypothesis test on the distance D: Test I

We can approximate the distribution of D with the normal distribution, i.e $D \sim N_1(\mu_D, \sigma_D^2)$. μ_D is the unknown expected value of D. We have the following hypothesis test:

$$\begin{aligned}H_0 &\;:\; \mu_D = 0\\ H_1 &\;:\; \mu_D > 0\end{aligned}$$

Under H_0 we have that $D \sim N(0, \sigma_D^2)$. Even though σ_D is unknown, we can estimate it from (6.2). We reject H_0 on significance level α if $\frac{D_0}{\hat{\sigma}_D} > k_{\alpha,1}$, where $k_{\alpha,1}$ is the $100(1-\alpha)$-quantile of the standard normal distribution. With $\alpha = \frac{1}{500}$, we have $k_{\alpha,1} = 2.878$. The separation factor ω is connected to the hypothesis test as follows:

$$\omega = \frac{D}{k_{\alpha,1}\sigma_D} \tag{6.3}$$

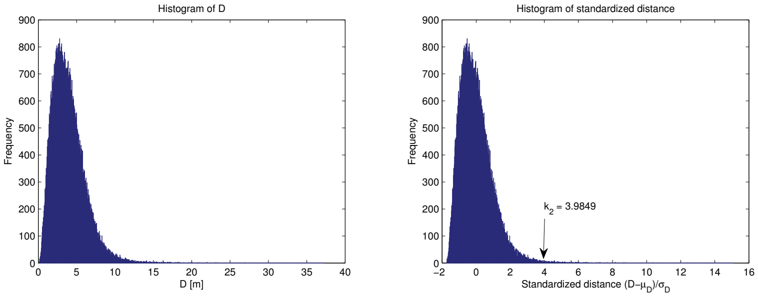


Figure 6.3: Histogram of $D \sim f_D(d)$ with $E(D) = \mu_D$ and $\text{Var}(D) = \sigma_D^2$ from a typical simulation (left). Histogram of $\frac{D-\mu_D}{\sigma_D}$ (right). The empirical 99.5%-quantile k_2 is found from the distribution of the standardized distance in hypothesis test II.

Rejection of H_0 is equivalent of calculating a separation factor $\omega > 1$. This test can easily be extended when we introduce the wellbore diameters. Let d_u and d_v be the wellbore diameters at u and v, respectively. The hypothesis test is then:

$$\begin{aligned} H_0 &\quad : \quad \mu_D \leq \frac{d_u + d_v}{2} \\ H_1 &\quad : \quad \mu_D > \frac{d_u + d_v}{2} \end{aligned}$$

In the same matter, we can extend the separation factor of (6.3):

$$\omega_1 = \frac{D - \frac{d_u+d_v}{2}}{k_{\alpha,1}\sigma_D} \tag{6.4}$$

The normality assumption on the distribution of D can be discussed. A normal distribution is defined on $\mathbb{R}$, while D only can have non-negative values. We have discussed the normality assumption in Appendix C. The appendix indicates that the a normal distribution is a good approximation to the distribution of D, especially for large distances D. "Large distance" is relative to the uncertainty of the distance, so that there is a very small risk of well-collision, which usually is the case.

6.4.2 Hypothesis test on the empirical distance D: Test II

We assumed that the distribution of D was normally distributed in Section 6.4.1. This was based on normally distributed position uncertainties in the wells. Let now $u \sim f_U(u)$ and $v \sim f_V(v)$, where $f_U(u)$ and $f_V(v)$ are general statistical distributions, which we are to simulate from. We can then use Monte Carlo simulations to simulate $D \sim f_D(d)$, since $D = \sqrt{(u-v)^T(u-v)}$. The difference between this test and the test in Section 6.4.1 is that $f_D(d)$ is now a general statistical distribution, which we are capable of simulating from. We denote $E(D) = \mu_D$ and $\mathrm{Var}(D) = \sigma_D^2$, and can use the same hypothesis test as in Section 6.4.1:

$$\begin{aligned} H_0 &: \quad \mu_D \leq \frac{d_u + d_v}{2} \\ H_1 &: \quad \mu_D > \frac{d_u + d_v}{2} \end{aligned}$$

Let $D_1, .., D_n$ denote n independent realizations from $f_D(d)$ simulated by Monte Carlo simulations. We look at the transformation:

$$Z_i = \frac{D_i - \frac{d_u+d_v}{2}}{\sigma_D}$$

We have that $E(Z_i) = \frac{\mu_D - \frac{d_u+d_v}{2}}{\sigma_D}$, and $\mathrm{Var}(Z_i) = 1$. Under H_0, we have that $E(Z_i) = 0$. If n is large, $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} D_i \approx \mu_D$. We can calculate $k_{\alpha,2}$ from the empirical distribution of $W_i = Z_i - \frac{\bar{d} - \frac{d_u+d_v}{2}}{\sigma_D}$, as $E(W) \approx 0$ and $\mathrm{Var}(W) = 1$. Figure 6.3 shows an example of how $k_{\alpha,2}$ is estimated from the simulations. The separation factor is calculated as in (6.4), and we reject H_0 if

$$\omega_2 = \frac{D_0 - \frac{d_u+d_v}{2}}{k_{\alpha,2}\hat{\sigma}_D} > 1$$

6.5 The probability of collision

The hypothesis tests in Section 6.4 allow us to control the probability of accepting a wellplan given that the wellplan results in a well-collision. If $\omega > 1$, we are $100(1-\alpha)\%$ certain that the expected value of D, μ_D, is greater than the sum of the radii, where α is the significance level of the test. The conclusion is based on the assumption that the computed centre-to-centre distance from the minimum curvature formulas is an unbiased estimate of μ_D. This decision criterion is extensively used in the industry today ([5] and [26]). A second alternative in anti-collision calculations is to compute the probability of collision, p_c. We can calculate p_c from either the statistical distribution of the difference vector r or from the statistical distribution of the separation distance D. Let $r \sim f_r(r)$ and $D \sim f_D(d)$. p_c can be computed from $f_r(r)$ as follows:

$$p_c = \int_{||r|| \leq \frac{d_u+d_v}{2}} f_r(r)\mathrm{d}r \tag{6.5}$$

(6.5) is a three-dimensional integral which usually is impossible to solve analytically, even for the multivariate normal distribution. The computation of p_c from $f_D(d)$ has a simpler expression:

$$p_c = P(D \leq \frac{d_u + d_v}{2}) = \int_0^{\frac{d_u+d_v}{2}} f_D(x)\mathrm{d}x \tag{6.6}$$

It is important to know the difference between the probability of collision, p_c, and the significance level which is chosen in the hypothesis tests, as they are two different risk interpretations in anti-collision theory. Neither (6.6) is possible to solve analytically, since we lack the analytical expression of $f_D(d)$. But both (6.5) and (6.6) are possible to solve numerically if we know the statistical distributions. If we do not know the analytical expressions of the statistical distributions, we can always estimate p_c from simulations by Monte Carlo integration. But p_c is very small even for wells that are designed to crash, especially if the position uncertainties are quite large. This means that Monte Carlo integration, even for a large number of realizations, gives a poor estimate of the probability of collision. Variance reduction techniques, as importance sampling [19], may give better estimates of the probability of collision. But the Monte Carlo estimates of the probability of collision are often very small ($\sim 10^{-5}$) even for wells that are designed to crash. The hypothesis tests, which lead to the separation factors, are intuitive methods which perform statistical tests on the wellplan. If the probability of collision, as it is presented in (6.6), were to be used in anti-collision calculations, one would have to have a logical interpretation of what the maximum tolerance probability would be. It is much more intuitively interpreted through a hypothesis test where you control the probability of accepting a wellplan given that it results in a well-collision.

Chapter 7

Results of the anti-collision calculations

7.1 Introduction

In Chapter 7 we present the results the anti-collision calculations. We focus on comparing the separation factor from the NIG error model against the separation factor from the normal error model. We restrict ourselves to wells of equal diameters $d_u = d_v = 8.5$ inches. Our choice of diameter is not critical with respect to the conclusions derived from the results. In the cases where we use synthetic wells, we restrict ourselves to straight well-geometries. It is probably easier to interpret the results from straight wells than from complex wells. We set the significance level to $\alpha = 0.002$. Section 7.2 explains how we plan to present the anti-collision calculations. Since we deal with a three-dimensional distribution, we have challenges in visualizing the results in a proper manner. An illustrative anti-collision figure is also presented in Section 7.2. Section 7.3-7.5 show three anti-collision cases which we analyze thoroughly in this thesis. Section 7.6 summarizes some of the general results which we have obtained during the anti-collision calculations.

7.2 Presentation of anti-collision simulations

In Section 5.4 we presented a figure which summarized the position uncertainty simulation, with Figure 5.3 as an example. Our anti-collision calculations are also simulation based, and we need to present the simulations in a similar figure as Figure 5.3. In every result, we present three different separation factors. The separation factors are presented in Table 7.1. ”Normal calculated” is the calculations which are common in the oil- and gas industry today, using the

Case	Separation factor	ε_{16}	$f_D(d)$	$\hat{\sigma}_D$
Normal calculated	ω_1	Normal	Normal	Normal
Normal simulated	ω_2	Normal	Empirical	Empirical
NIG simulated	ω_3	NIG	Empirical	Empirical

Table 7.1: Overview of the different separation factors that are analyzed in the anti-collision calculations.

normal error model. ω_1 is calculated from hypothesis test I (Section 6.4.1). "Normal simulated" is also based on the normal error model, but we use empirical distributions of D to compute ω_2, as in hypothesis test II (Section 6.4.2). "NIG simulated" is based on the NIG error model. The computation of ω_3 is also based on hypothesis test II. The only difference between the separation factors is how $\hat{\sigma}_D$ and k_α are calculated. When we compare the separation factors, we use the term "conservative". If $\omega_3 < \omega_1$, ω_3 is considered more conservative than ω_1 with respect to the conclusions that are made based on the separation factors. The general formula for calculating the separation factor is:

$$\omega_i = \frac{D_0 - \frac{d_u + d_v}{2}}{k_i \hat{\sigma}_{D,i}} \quad i = 1, 2, 3 \quad , \tag{7.1}$$

where D_0 is the calculated minimum centre to centre distance, k_i is the desired quantile of the distribution of D under H_0 and $\hat{\sigma}_{D,i}$ is the estimate of the standard deviation of D.

Figure 7.1 is an example of how we present results of anti-collision calculations. The three subplots in the right column in Figure 7.1 show the wellpaths projected onto the NE-, NV- and EV-plane. The reference well is plotted in blue, and the offset well is plotted in black. A position uncertainty simulation is performed on a given station in the reference well, and the 3D Closest approach scanning method is used for finding the closest point in the offset well. A similar position uncertainty simulation is performed in the closest point in the offset well. Both simulations are projected onto the NE-, NV- and EV-plane and the N-, E- and V-axis, as in Figure 5.3. If both simulations overlap in every projection, there may be a high risk of collision. The risk is considered significant under a given model, as in Table 7.1, if $\omega < 1$. The upper subplot in the second left column shows the simulated distribution $f_D(d)$. The middle subplot in the second right column is unchanged since Figure 5.3, and shows the NIG-parameters that are chosen on the declination error source.

The upper subplot in the second right column displays the calculations for the cases explained in Table 7.1. "Normal calculated centre to centre distance" is the minimum distance between the reference- and offset well, calculated from the minimum curvature formulas. The "Normal calculated centre to centre st.dev." is the estimate of σ_D based on the error model in Section 4.2, which is calculated from (6.2). These two estimates are directly comparable with the simulations results displayed beneath. Figure 7.1 shows two synthetic wells with high risk of colliding. We see that the normal calculated centre to centre distance differs

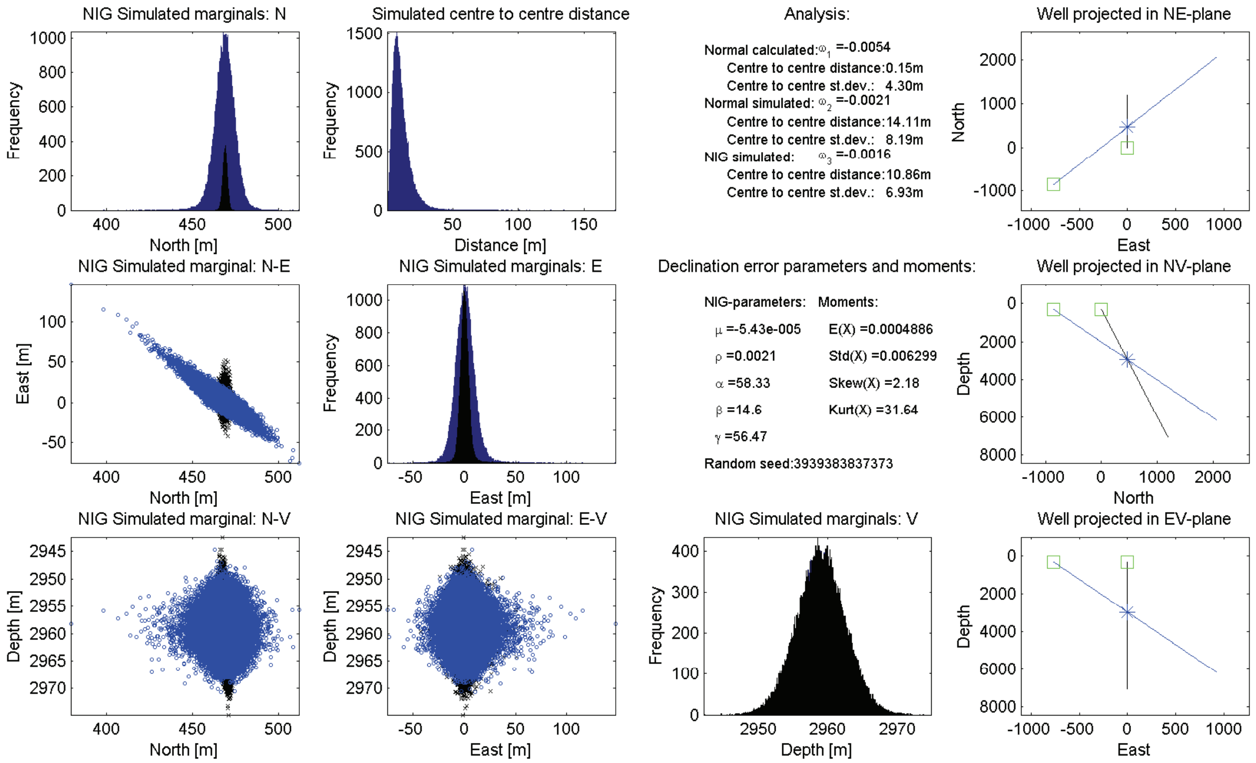


Figure 7.1: An example of an anti-collision calculation based on simulations. This simulation shows two synthetic wells that have a high risk of intersecting. All separation factors are much lower than 1, which indicates that this reference well would never be drilled after this plan.

much from the simulated distances. The reason why they differ is that the normal approximation of $f_D(d)$ is poor for so small distances relative to the uncertainty, as we discussed in Section 6.4.

We focus on analyzing the difference between ω_1 and ω_3 given that the wells are separated in the NE-plane. As we saw in Section 4.2, the declination error only affects the azimuth angle, which leads to differences between the error models in the NE-plane. Given a candidate point in the offset well, we can shift the reference well on a grid in the NE-plane, and perform the anti-collision calculations as in Figure 7.1 for each point in the grid. Then we are able to analyze how the separation factor changes in the NE-plane. Since the declination error only affects the NE-coordinates, we see no need for analyzing the difference between separation factors along the V-axis. We will connect the results to the directions of the empirical eigenvectors of the simulations.

7.3 Anti-collision Case I

Reference well	845	36.13	30.65
N	0	0.707	0.707
E	-1	0	0
V	0	0.707	-0.707

Offset well	21.65	15.27	1.38
N	0.166	-0.148	0.975
E	0.984	0.095	-0.153
V	-0.070	0.984	0.162

Table 7.2: The empirical eigenvalues (top) and eigenvectors (below) for the reference well and offset well of anti-collision case I.

We analyze the case where the offset well is a real well and the reference well is a synthetic well. The reference well is a straight well with constant inclination angle $I = 45$ and azimuth angle $A = 0$. In this case we look at the collision risk in a given point in the offset well ($D = 3200, I = 7.5, A = 274$) when we shift the reference well in different positions in the NE-plane. Figure 7.2 shows this case where the reference well is calculated to be separated from the offset well with 80 meters to the east. The reference well is directed to the north, and has most uncertainty along the E-coordinate. The offset well is almost directed to the west at this point, but it has been directed southwards almost up to this point. Therefore, the uncertainty is largest along the E-coordinate. Table 7.2 shows the empirical eigenvectors and eigenvalues of anti-collision case I. The first eigenvector of the reference well (eigenvalue 845) is dominating in the analyses, which is also easily seen from Figure 7.2.

We see that $\omega_1 > 1$ in Figure 7.2, indicating that there is no significant risk of well-collision given that the normal error model is the true model. But $\omega_3 < 1$

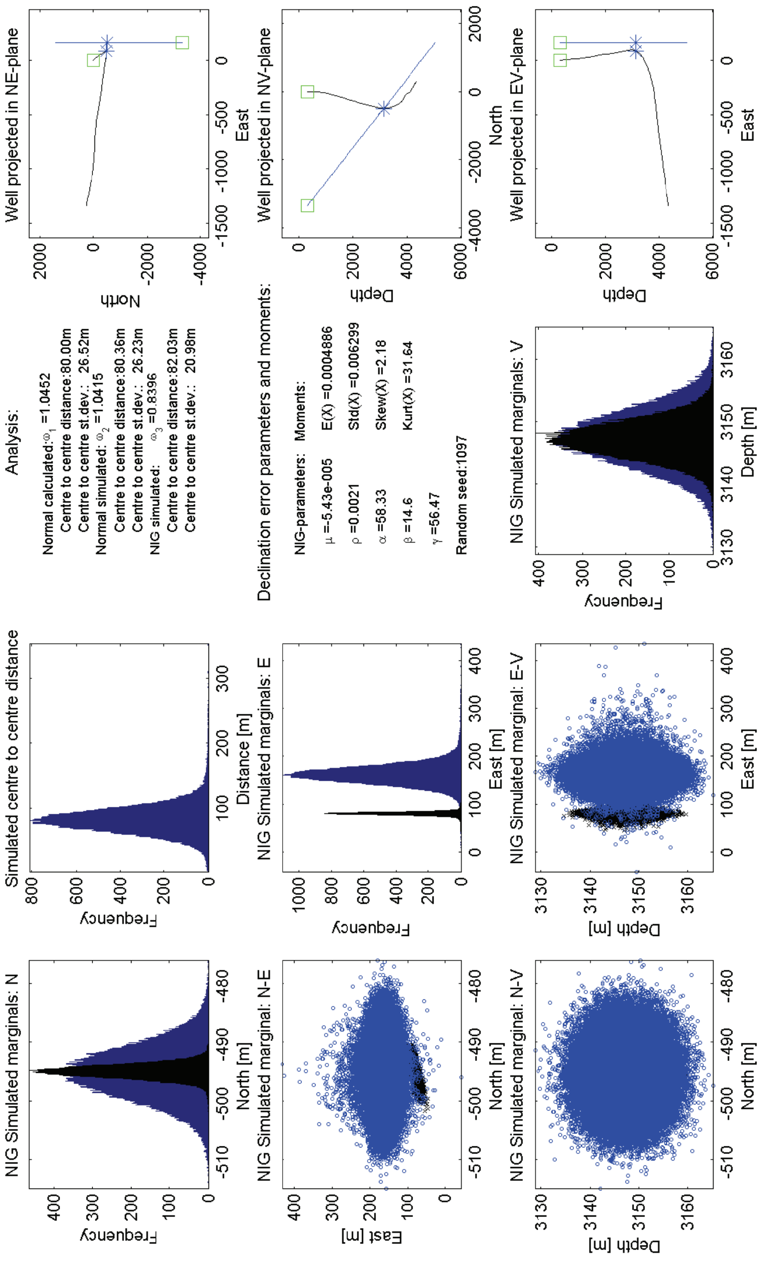


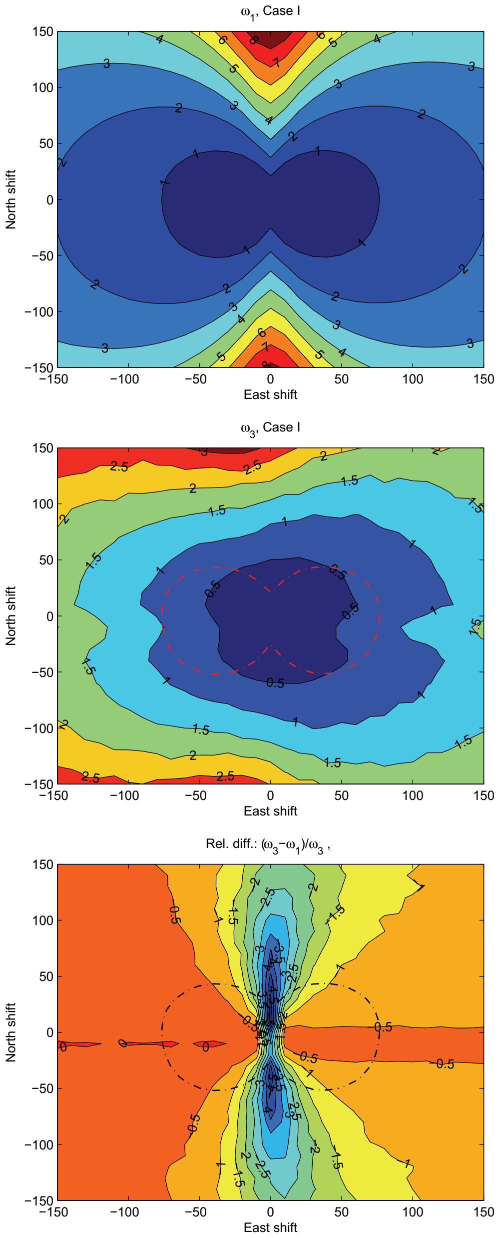
Figure 7.2: Anti-collision case I. A synthetic northern directed well with constant inclination ($I = 45$) is close to intersecting with a real offset well. We see that $\omega_1 > 1$, signalizing that there is no significant risk of well-collision given that the normal error model is correct, while $\omega_3 < 1$, signalizing the opposite for the NIG error model.

indicates the opposite conclusion given that the NIG error model is the true model. $\hat{\sigma}_{D,3}$ is smaller than $\hat{\sigma}_{D,1}$ which is expected since the NIG-distribution has a larger mode than the corresponding normal distribution. We also see that $\hat{\mu}_{D,3} = 82.03$, which indicates that there is 2 meter bias between the minimum centre to centre distance and the simulated distributions. The difference can be explained by the deviation in expected position, as discussed in Section 5.7. $\hat{\mu}_{D,2} = 80.36$ indicates that there is 36 cm difference between the calculated and simulated centre-to-centre distance. This small difference is most likely caused by random effects from the simulation.

We perform the same type of analysis as shown in Figure 7.2 several times, but now we shift the reference well on a equidistant grid in the NE-plane. The offset well is kept in the same position the whole time. In this way, we can control the separation of the wells in the NE-plane. With this type of analysis we can see what directions that yield different results between the two error models. Figure 7.3 shows the analyses. We see from Figure 7.3 that there are large differences between the two separation factors ω_1 (upper) and ω_3 (centre). We see that the difference is largest if we shift the reference well along the north coordinate. ω_1 is symmetrical with respect to the shift direction, while this is not the case for ω_3. This is explained by the skewness of the NIG error model. ω_1 has a distinct cusp [24] shape along shifts in the north coordinate, which we will see can be explained by the normal approximation.

The contour plots in Figure 7.3 can be further explained by the corresponding estimates of $\hat{\sigma}_D$ and the quantile value k_3. Figure 7.4 shows contour plots of these estimates. Recall that the quantile value for the normal error model is $k_1 = 2.878$. We see that k_3 is greater than k_1 everywhere in our analyses. This is explained from the NIG-distribution which makes the distribution $f_D(d)$ more heavy-tailed, as shown in Figure 6.3. The contour plot of k_3 and $\hat{\sigma}_{D,3}$ are non-symmetric, as ω_3, with higher values for eastern shifts than for western shifts. k_3 is largest for shifts along the N-coordinate. Both wells have heavy tails along the E-coordinate. A shift in the N-coordinate results in a more heavy-tailed distribution $f_D(d)$ than in the E-coordinate, which results in a larger quantile k_3.

Figure 7.5 shows the contour plots of the normal simulated ω_2, and the relative differences $\frac{\omega_2-\omega_1}{\omega_2}$ and $\frac{\omega_2-\omega_3}{\omega_2}$. From these contour plots we see that the normal approximation in ω_1 makes a large difference along the N-axis. The relative difference between ω_2 and ω_3 is not so big as between ω_1 and ω_3 in Figure 7.3. The cusp in the contour plot of ω_1 causes large differences between ω_1 and ω_3, but the cusp is not apparent in the contour of ω_2. The cusp is a direct effect of the normal approximation of $f_D(d)$. We see that $\frac{\omega_2-\omega_3}{\omega_2}$ is smallest for shifts along the E-axis, which is the axis of largest uncertainty (Largest eigenvector in Table 7.2). It is also worth noticing that $\frac{\omega_2-\omega_3}{\omega_2} > 0$, which implies that $\omega_2 > \omega_3$. ω_3 is then the most conservative separation factor of the ones that are calculated. Regarding the normal approximation, we see that $\frac{\omega_2-\omega_1}{\omega_2} < 0$, indicating that ω_2 is more conservative than ω_1. This indicates that occasions may occur when a wellplan is accepted ($\omega_1 > 1$), but the real distribution of D



ω_1, Case I

ω_3, Case I

Rel. diff.: $(\omega_3-\omega_1)/\omega_3$,

Figure 7.3: A thorough analysis of the anti-collision case I, where the straight reference well ($I = 45, A = 0$) is shifted on a grid in the NE-plane and the anti-collision calculations are performed against the point ($D = 3200, I = 7.5, A = 274$) in the offset well. Contour plot of ω_1 is shown in the upper figure, ω_3 is shown in the middle figure, and the relative difference $\frac{\omega_3-\omega_1}{\omega_3}$ is shown in the lower figure. We see that there are large differences between ω_1 and ω_3, where ω_3 is the most conservative separation factor with respect to the hypothesis tests.

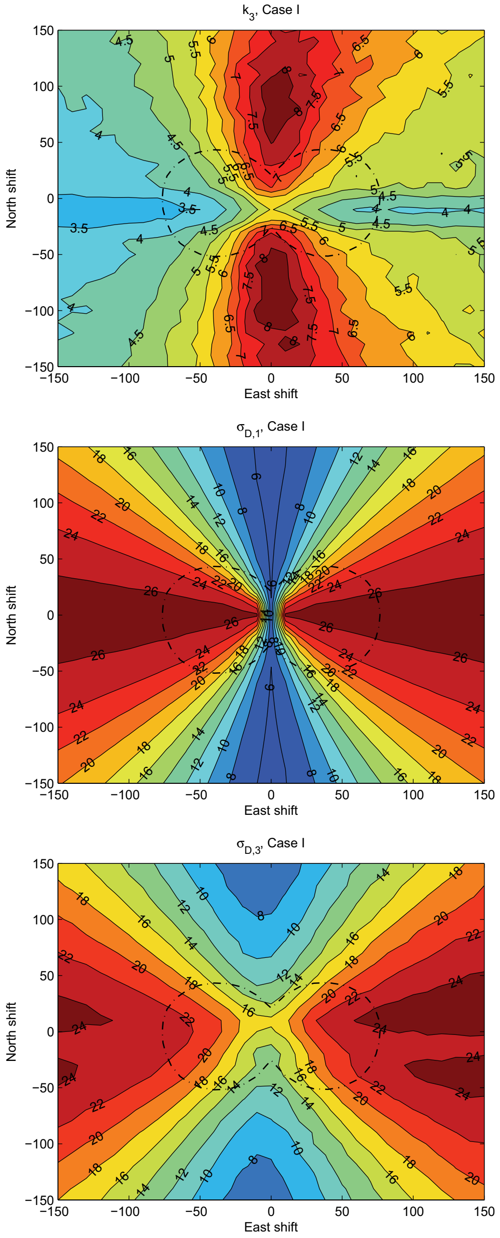


Figure 7.4: Contour plots of the empirical quantile k_3 (upper), the empirical estimate $\hat{\sigma}_{D,1}$ (middle) and $\hat{\sigma}_{D,3}$ (lower) for the anti-collision case I where the reference well is shifted on a grid in the NE-plane. The dashed line indicates where $\omega_1 = 1$.

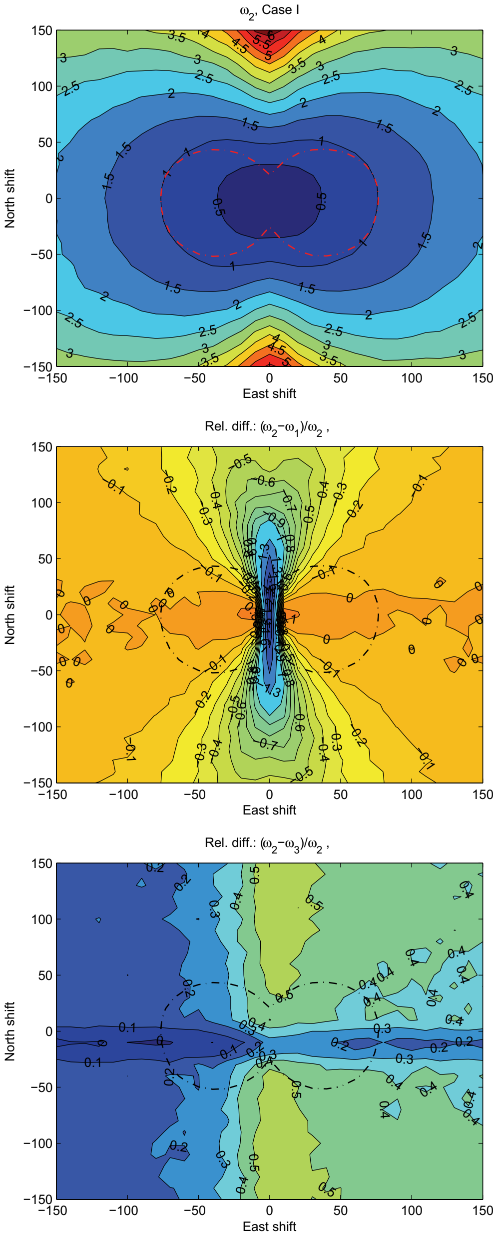


Figure 7.5: Contour plots of the normal simulated ω_2 (upper), and the relative difference $\frac{\omega_2-\omega_1}{\omega_2}$ (centre), and the relative difference $\frac{\omega_2-\omega_3}{\omega_2}$ (lower) for anti-collision case I.

is so different from the normal distribution that $\omega_2 < 1$.

7.4 Anti-collision Case II

Both wells	172	38	17
N	-0.980	0.080	-0.182
E	-0.199	-0.356	0.913
V	0.009	0.931	0.365

Table 7.3: The empirical eigenvalues (top) and eigenvectors (below) of anti-collision case II.

Anti-collision case II is a case where we look at two wells that are identical and parallel. The case that the wells are identical is not a realistic scenario, but for the purpose of this thesis we are able to study the effects of having identical position uncertainties. We perform our analysis in the point $(D = 5000, I \approx 66, A \approx 305)$ in the offset well, and shift the other well in a grid in the NE-plane, as we did in case I. Figure 7.6 shows the position uncertainty at the approximate position which we perform our analyses. We see that the N-coordinate is the axis with the largest uncertainty. This is intuitive since the well has travelled most in the western direction relatively to the starting position. Table 7.3 shows the empirical eigenvalues and eigenvectors of the well. The first principal axis is approximately directed N-S. The second principal axis is directed most along the V-axis, while the third principal axis is directed most along the E-axis. We have used the same position uncertainty earlier in the thesis.

We shift one of the wells in a equidistant grid in the NE-plane and compute the separation factors for each grid-node, as we did in case I. Figure 7.7 shows contour plots of ω_1, ω_3 and $\frac{\omega_3 - \omega_1}{\omega_3}$. In case I we saw that $\omega_3 < \omega_1$ in the whole grid. In most of the contour plot of Figure 7.7 we see that $\omega_3 > \omega_1$, but the most interesting parts of the plot is around the contour where $\omega_1 = 1$, where the decisions are made by today's standards. Here we observe that we have shift directions where $\omega_3 < \omega_1$ and other shift directions where $\omega_3 > \omega_1$. We see that $\omega_1 > \omega_3$ along the third principal axis, and $\omega_1 < \omega_3$ along the first principal axis. We are not able to see the effects along the second principal axis since it is directed mostly along the V-axis. We see the same symmetry in the contour plot of ω_1 in Figure 7.7 as we did in Figure 7.3 for case I. The cusp which we saw in case I is seen in the contour of ω_1 along the third principal axis. A shift along the first principal axis gives the largest change in ω_1. The same effect is seen for ω_3 too, but the change in ω_3 is not that big as for ω_1.

Figure 7.8 shows contour plots of k_3, $\sigma_{D,1}$ and $\sigma_{D,3}$. In case I we saw that $\omega_3 < \omega_1$, which indicates that $k_3\sigma_{D,3} > k_1\sigma_{D,1}$. In the areas where $\omega_3 > \omega_1$ in case II we have that $\sigma_{D,3}$ is much smaller than $\sigma_{D,1}$, and it decays faster with increased separation distance. The difference is largest with a shift along

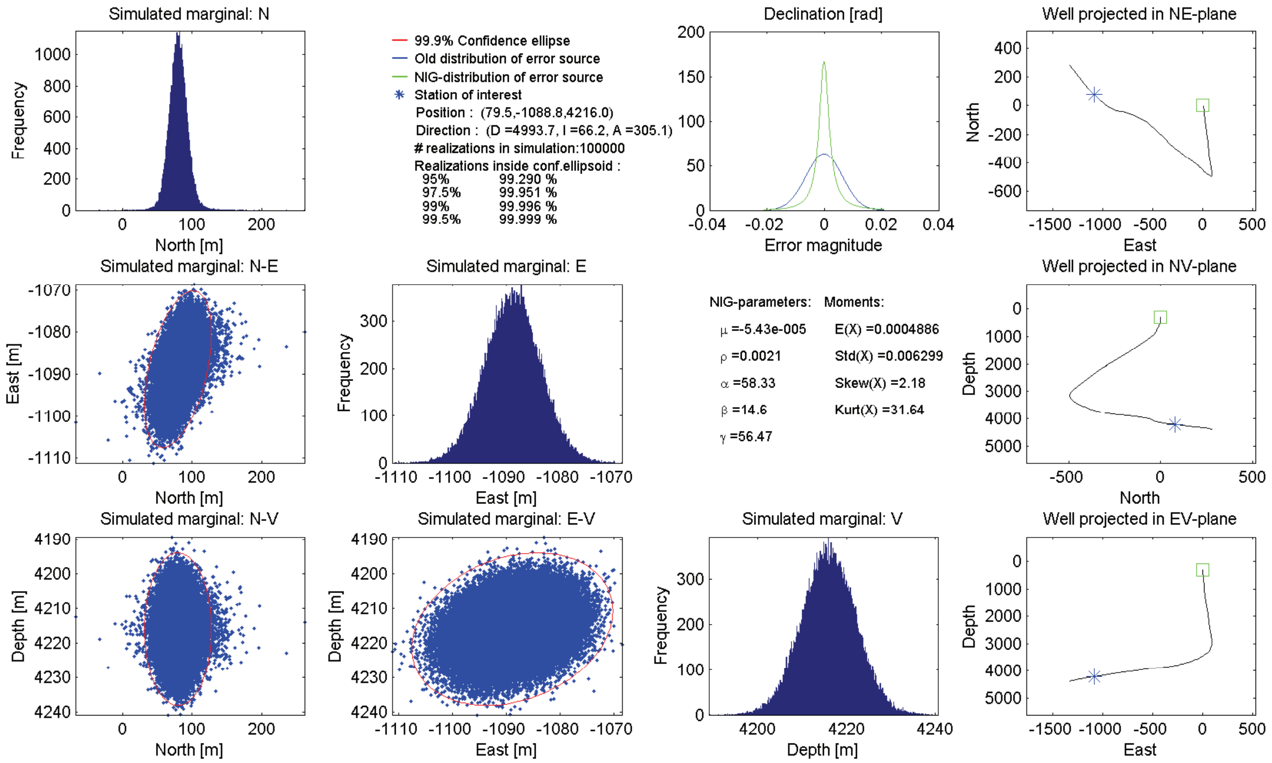


Figure 7.6: Anti-collision case II. We perform the anti-collision analyses of two identical parallel wells at $D = 5000$ meters measured depth. The inclination angle is approximately 66 degrees, and the azimuth angle is approximately 305 degrees. Both wells have identical position uncertainty because of the equal well geometries, so only one of them are shown in this figure.

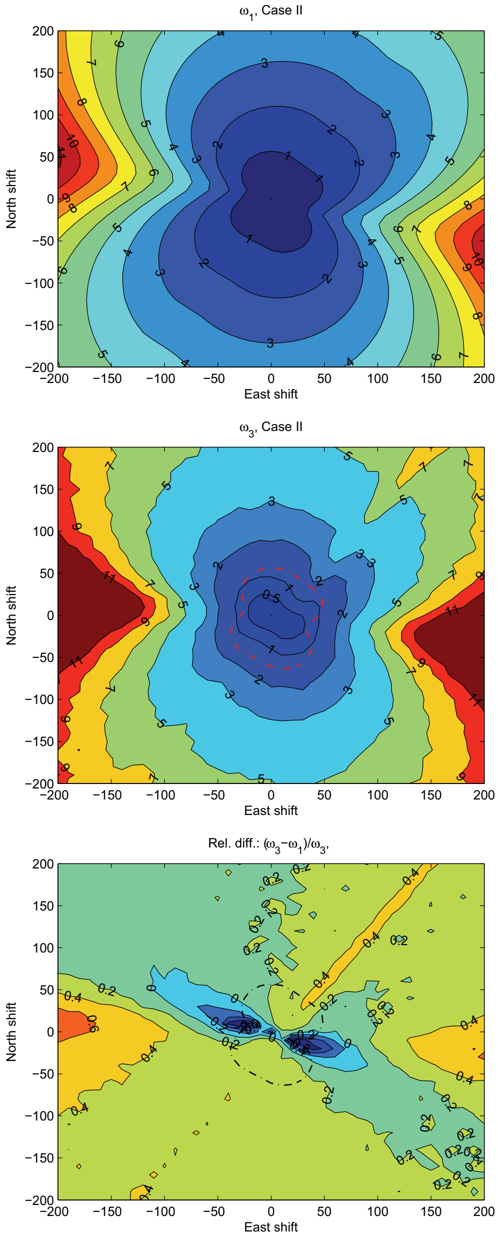


Figure 7.7: A thorough analysis of the anti-collision case II where one of the wells are shifted in a grid in the NE-plane. Contour plot of ω_1 (upper), ω_3 (middle), and the relative difference $\frac{\omega_3-\omega_1}{\omega_3}$ (lower). We see that the relative difference is both greater and smaller than 0, which gives no certain answer for which error model is the most conservative for this case.

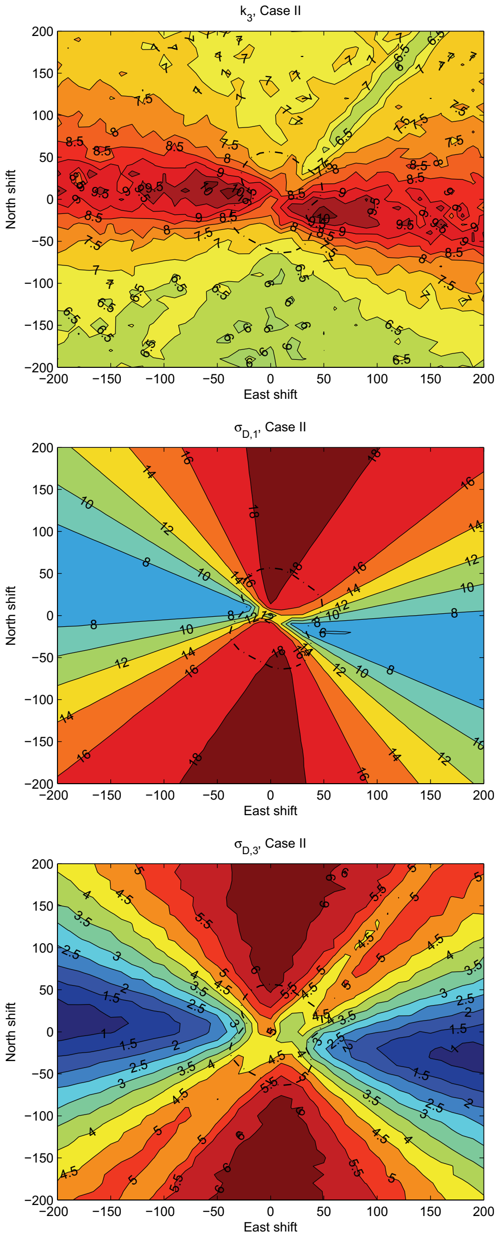


Figure 7.8: Contour plots of the empirical quantile k_1 (upper), the empirical estimate $\hat{\sigma}_{D,1}$ (middle) and $\hat{\sigma}_{D,3}$ (lower) for the anti-collision case II where one of the wells are shifted in a grid in the NE-plane. The dashed line indicates where $\omega_1 = 1$.

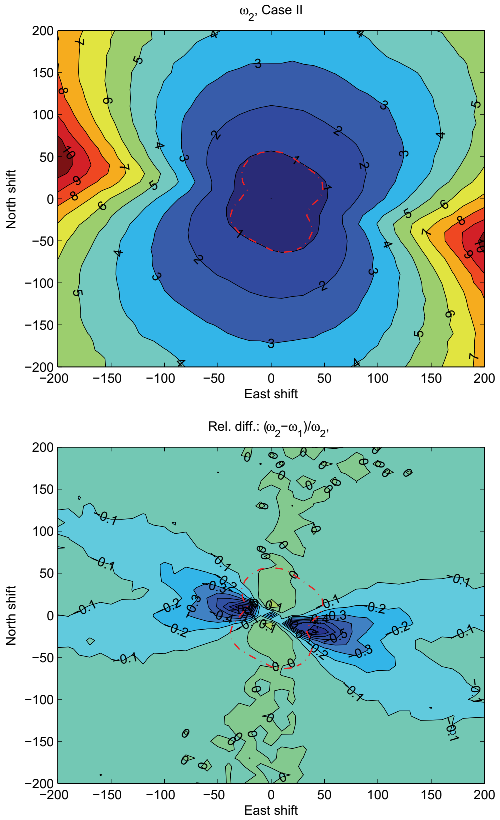


Figure 7.9: Contour plots of the normal simulated ω_2 (upper), and the relative difference $\frac{\omega_2-\omega_1}{\omega_2}$ (lower) for anti-collision case II.

the first principal axis. Both $\sigma_{D,1}$ and $\sigma_{D,3}$ show that a shift along the third principal axis gives the largest uncertainty in D.

Figure 7.9 shows the contour plot of the normal simulated ω_2 and the relative difference $\frac{\omega_2-\omega_1}{\omega_2}$. The normal approximation causes deviations between ω_1 and ω_2 along the first principal axis with $\omega_1 > \omega_2$, but along the third principal axis the normal approximation of $f_D(d)$ seem to be good. As in case I, we see that $\frac{\omega_2-\omega_1}{\omega_2} < 0$, making ω_2 more conservative than ω_1.

When we shift one of the identical position uncertainty distributions in case II, the distributions in the wells are parallel. A sample from a NIG-distribution have a higher frequency of realizations close to the mean than a sample from a normal distribution. The same effect is seen in the position uncertainty distributions, as shown in Figure 5.5. Two normal distributions (with different mean and equal variances) which overlap relatively much, overlap more than two NIG distributions with the same mean and variance as the normal distributions. This is the reason why $\sigma_{D,3} < \sigma_{D,1}$ and $\omega_3 < \omega_1$ for most kinds of shift in case II.

7.5 Anti-collision Case III

Reference well	644	33.15	6.22
N	-0.708	-0.352	0.612
E	0.706	-0.355	0.612
V	0.002	0.866	0.5
Offset well	658	33.15	6.22
N	-0.708	-0.352	-0.612
E	-0.706	0.355	0.612
V	0.002	0.866	-0.5

Table 7.4: The empirical eigenvalues (top) and eigenvectors (below) for the reference well and offset well of anti-collision case III.

Anti-collision case I analyzed a point in the offset well where the well had been pointed southwards up to this point. The reference well was a well pointing northwards. This gave us largest uncertainty along the E-axis for both wells. In anti-collision case II we analyzed two identical well geometries, which result in identical principal axes in both wells. In anti-collision case III we analyze two wells that intersect orthogonally to eachother in the NE-plane. The offset well is a straight well with $A = 315$ and $I = 60$. The reference well is a straight well with $A = 45$ and $I = 60$. Figure 7.10 shows an example of this case. The calculations are performed at $D = 3000$ in the offset well. In Figure 7.10 the wells are calculated to be separated by 70.71 meters. It is worth noting that the estimated normal simulated centre-to-centre distance is over 5 meters different from the normal calculated. Table 7.4 shows the empirical eigenvectors

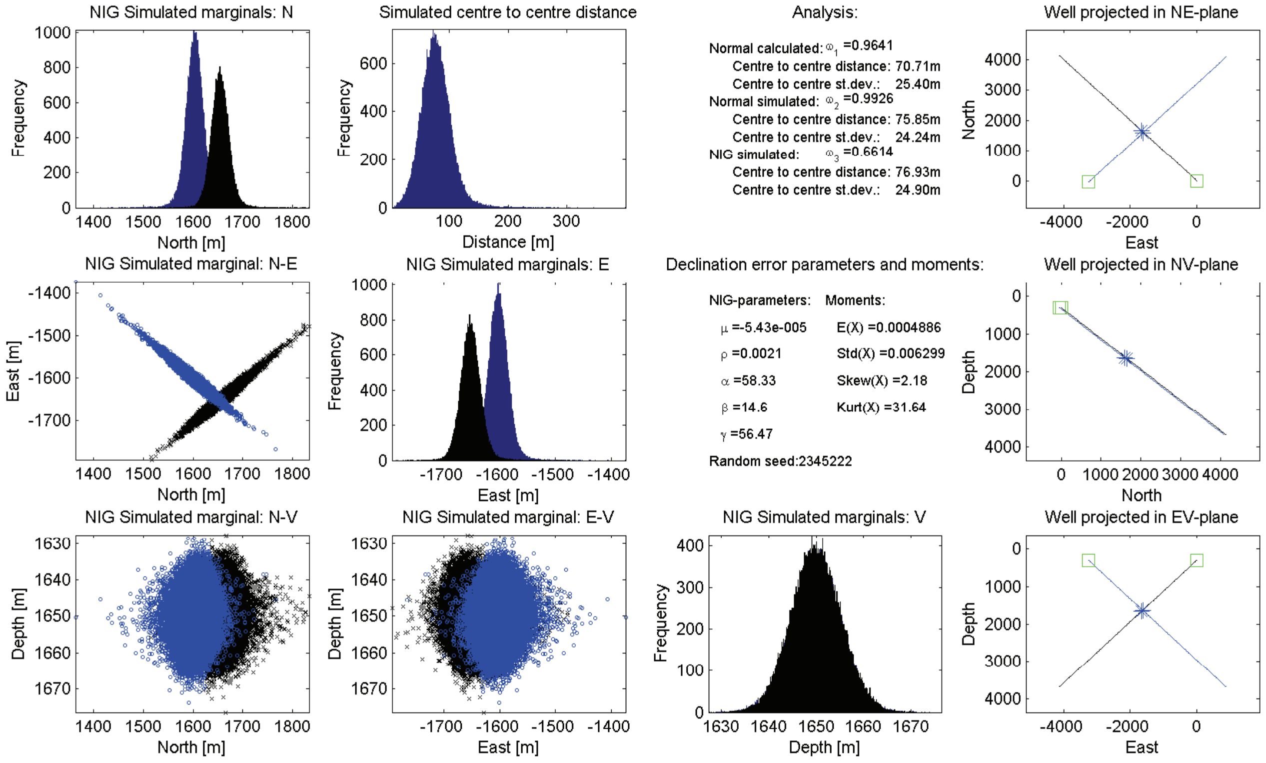


Figure 7.10: Anti-collision case III. The offset well is a straight well with $A = 315$ and $I = 60$. The reference well is a straight well with $A = 45$ and $I = 60$. We perform the anti-collision calculations at $D = 3000$ in the offset well. The calculated separation between the wells in this plot is 70.71 meters.

and eigenvalues of this case. We see that all eigenvectors of the offset well are orthogonal to the eigenvectors of the reference well in the NE-plane. The first principal axis of both wells indicate that the uncertainty is equally large along both the N- and E-axis. We shift the reference well in a grid in the NE-plane and analyze the differences between the separation factors ω_1 and ω_3.

Figure 7.11 shows the contour plots of ω_1, ω_3 and $\frac{\omega_3-\omega_1}{\omega_3}$ for anti-collision case III. The cusp which we saw in both case I and case II is not apparent in case III. This can only be explained by the largest eigenvectors which are orthogonal to eachother in the NE-plane. We see that $\omega_3 < \omega_1$ for the whole contour, making ω_3 the most conservative separation factor. In the relative difference contour $\frac{\omega_3-\omega_1}{\omega_3}$ we notice an area where the relative difference is large. Since both wells have $I = 60$, the uncertainty in depth becomes large, as we can see of the second eigenvectors in Table 7.4. The second principal axis is mostly directed along the V-axis. The large relative difference is then caused by the normal approximation which becomes poor along this principal axis. The cusps we saw in case I and II are now placed along this principal axis, and we observe the cross-section between the NE-plane and this axis.

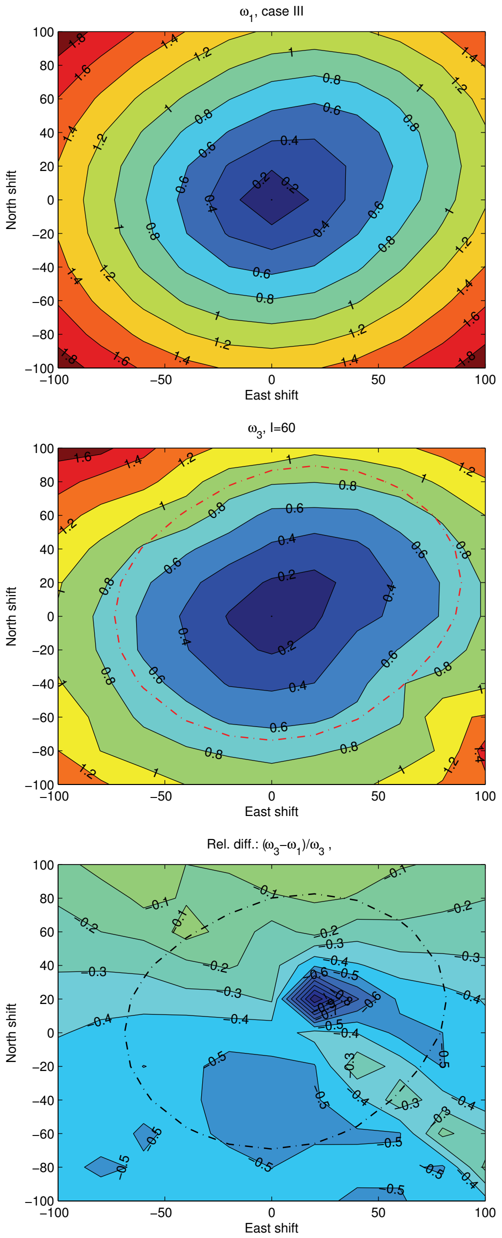


Figure 7.11: A thorough analysis of the anti-collision case III. Contour plot of ω_1 (upper), ω_3 (middle) and the relative difference $\frac{\omega_3 - \omega_1}{\omega_3}$ (lower).

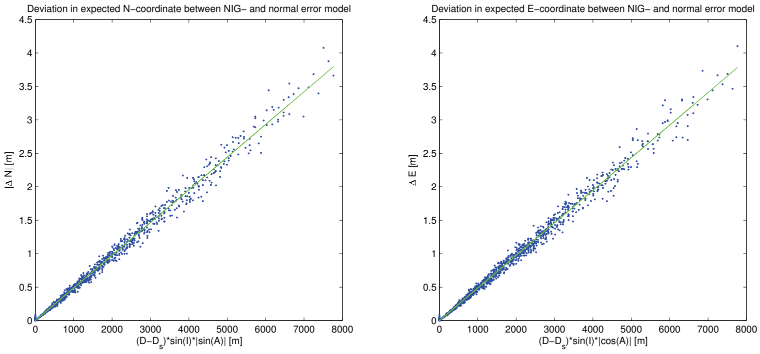


Figure 7.12: Simulation results of the deviation along the N- (left) and E-coordinate (right) between the expected position from the NIG error model and the calculated position from the minimum curvature formula. The simulation results are made on different measured depths from synthetic straight wells of different inclination angles and azimuth angles. The values along the x-axis are $(D_i - D_s)\sin I_i|\sin A_i|$ (left) and $(D_i - D_s)\sin I_i|\cos A_i|$ (right), where D_s is the measured depth at seabed, and D_i, I_i, A_i are the measured depth, inclination and azimuth of the wells. The green lines are fitted regression models.

7.6 General results

In Section 5.7 we discussed how a skewed declination error with $E(\varepsilon_{16}) \neq 0$ results in a deviation between the expected position in the NIG error model and the normal error model. Equation (4.29) allows us to compute the deviation in expected position between the NIG error model and the normal error model. We verify (4.29) by simulating the position from the NIG error model in synthetic straight wells of different inclination and azimuth. This is done on different measured depths in the straight wells. Figure 7.12 shows the simulation results, where we can see a significant trend in the deviations along both the N- and E-coordinate. Let ΔN be the deviation in expected N-coordinate, ΔE be the deviation in expected E-coordinate, and D_s be the measured depth at seabed where we assume zero uncertainty. The following predictors are estimated from the regression models of the simulation results:

$$\Delta\hat{N} = -4.89 \cdot 10^{-4}(D - D_s)\sin I \sin A$$
$$\Delta\hat{E} = 4.87 \cdot 10^{-4}(D - D_s)\sin I \cos A$$

We recall from Section 5.7 that $E(\varepsilon_{16}) = 4.88 \cdot 10^{-4}$, which coincide with the estimated regression parameters. The covariates of the regression model is chosen this way because of the analytical form of (4.29). As an example, let $D = 5000$ m, $D_s = 300$ m, $I = 30$, $A = 60$. We expect a deviation in expected position of $\Delta\hat{N} = -0.9942$ m, and $\Delta\hat{E} = 0.5721$ m, resulting in a total deviation of 1.1471

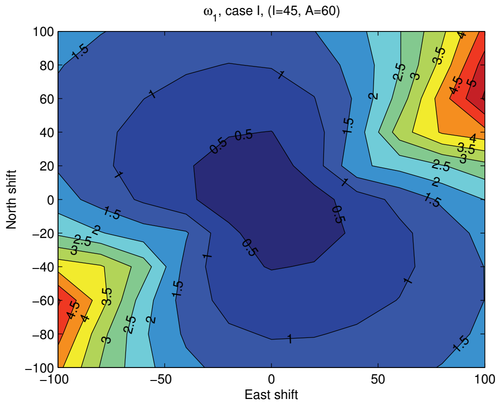


Figure 7.13: Contour plot of ω_1 for anti-collision case I, with the reference well having azimuth angle $A = 60$ instead of $A = 0$. By comparing with Figure 7.3 we see that our results look rotational invariant with respect to the azimuth angle of the reference well.

m in the NE-plane. For non-straight wells, one would have to compute the deviation in expected position from (4.29). Straight wells are considered to give the largest deviation. These results indicate that if $E(\varepsilon_{16}) \neq 0$ there will be a deviation in expected position. The deviation is usually not that big, but it may cause significant differences in the results, especially between the calculated and simulated centre to centre distances.

As we have seen, the position uncertainty from MWD magnetic directional surveying is highly dependent on the well geometry. The results are therefore highly dependent on the well geometries. We have chosen three illustrative cases which highlight this subject. Case I, in Section 7.3, is a typical example of two wells with different well geometries that collide. Even though it is not realistic to believe the the reference well is a straight well, the well is approximately straight in the section where there is a significant risk of well-collision. Case II, in Section 7.4, highlights what problems that might arise if we have parallel wells with approximately same geometry prior to the point of analysis. Figure 7.13 shows a contour plot of ω_1 when a straight reference well with $I = 45$ and $A = 60$ encounters the same point in the same offset well as in case I. We see the same effects occuring as in case I, except that the results are rotated. The results seem to be relatively rotational invariant with respect to the azimuth angle, although the numerical values are not the same. The cusps are slightly changing the shape, from being sharp when the eigenvectors are directed in the same direction in the NE-plane, to more curved when the eigenvectors are orthogonal to eachother in the NE-plane.

If we look at anti-collision case I, Figure 7.2, we see that the uncertainty in the reference well is greater than in the offset well. By changing the azimuth, we rotate the statistical distribution of the reference well. And since it is the largest distribution in magnitude, it ensures us that the results are rotated without any significant change. By looking at the eigenvector-directions of the reference well as we did in Section 7.3 from Table 7.2, we would easily could explain the results through the eigenvector-directions.

If we decrease the inclination angle I, the differences between the results decrease. This is because the declination error only affects the N- and E-coordinate with a factor proportional to $\sin I$. And if the reference well is close to vertical, the declination error has no effect on the total position uncertainty, and we are not able to see any significant difference between the two error models.

The normal approximation of $f_D(d)$ were discussed in Section 6.4.1 and Appendix C. The approximation is sufficiently good when there is a small risk of well-collision. Our results show large relative differences between ω_1 and ω_2 for small shifts along the largest principal axes. A shift along the largest principal axis is the direction where the risk of well-collision is reduced the least. We have seen that $\omega_1 > \omega_2$ especially along the largest principal axes. Given that the normal error model is the true error model, this may lead to situations where $\omega_1 > 1$ indicates that a wellplan can be built, while the true distribution of D would result in $\omega_2 < 1$ and the opposite conclusion.

Of our results we also observe that $\hat{\sigma}_{D,1} < \hat{\sigma}_{D,3}$ for all cases. A NIG-distribution has a larger area close to the mean compared to a normal distribution, and simulations of one well of the NIG error model give more dense realizations near the expected value than simulations of the normal error model. This results in a smaller variation of the distance between points in each simulation. The NIG-distribution has heavier tails than the normal distribution. This results in a heavier tail on the distribution of D, which results in a larger quantile value which is used in computing the separation factor ω.

Chapter 8

Closing remarks

8.1 Conclusions

There are obvious differences in the position uncertainty between the chosen normal error model and the chosen NIG error model. The declination error is the only difference between the error models. The declination error affects only the azimuth angle A, which again affects only the uncertainty in the NE-plane. There is no difference between the error models along the V-axis. A skewed and heavy-tailed statistical distribution on the declination error results in a skewed and heavy-tailed statistical distribution on the position uncertainty. A heavy-tailed declination error will in general affect all three principal axes of the position uncertainty. Our analyses do not conclude that the NIG error model have principal axes directed in other directions than the normal error model. The skewness of the position uncertainty is very dependent on the well geometry. When the declination error model has a small non-zero expectation there is a small deviation between the expected positions of the two error models.

Our results show that the use of the NIG error model usually gives a more conservative result than the use of a normal error model. The exceptions are seen for wells with similar geometries. When both the reference well and the offset well have equal geometries, they have equal position uncertainties. In these cases the normal error model is more conservative than the NIG error model. It is very hard to derive any general conclusion from our results, as the position uncertainties are very dependent on the well geometry.

By approximating the distribution of the minimum centre to centre distance D with a normal distribution, there may occur situations where a wellplan is accepted when the true distribution of D would result in the conclusion that the wellplan should not be accepted. These situations are most likely to occur when two wells are separated along an axis of large uncertainty, i.e. σ_D is large relative to the calculated minimum centre to centre distance D_0. Our results show a maximum relative difference in separation factor of 2 between the normal calculated separation factor and the normal simulated separation factor. In the same cases, the relative difference between the normal calculated- and the NIG simulated separation factor is even higher than 2. This results in a higher risk of well-collision than the chosen maximum tolerance probability when the distribution of D deviates from a normal distribution.

Given that the chosen NIG-distribution for the declination error is the true distribution, it may be wise to consider other approaches for anti-collision calculations than the normal theory which is currently practised. If two wells are separated in certain directions we have seen that there may be large differences in the results between the two error models. The normal approximation that is currently used in the hypothesis test of the anti-collision calculations becomes poor when the separation factor is close to 1, i.e. in the situations where we start to have a significant risk of well-collision. The arguments for the use of the current method is that it is easy to understand, easy to calculate analytically, and the statistical distribution used under the null hypothesis is analytically tractable. The use of an empirical distribution in the hypothesis test, which we are forced to use under the NIG error model, leads to quantile-values which change between every analysis. The simulation approach is also more time-consuming than the analytical approach, without this being a significant argument for the current method. An argument against the use of the current method is that it may cause us to accept a wellplan on a given significance level when the true probability of type I error is larger than the significance level. But as we have seen in this thesis, the normal error model is in some cases the most conservative method.

Corrective actions on the existing method may be an easier alternative than changing the declination error distribution. Our results indicate that the most important change would be to simulate the distribution of the minimum centre to centre distance, i.e. calculate the position uncertainties from the normal error model, and simulate the centre to centre distance distribution $f_D(d)$. An alternative to simulation of $f_D(d)$ would be to analytically compute $f_D(d)$ given the position uncertainties. Even though simulation of $f_D(d)$ does not correct for the heavy-tailed behaviour of the declination error, we have seen that the results are improved by avoiding the normal approximation of $f_D(d)$. The differences between the normal simulated- and the NIG simulated results are smaller than the difference between the normal calculated- and the NIG simulated results. If the declination error is believed to be non-central, it is possible to correct for by introducing a bias in the expected position. This bias can easily be calculated from the well geometry and the expected value of the declination error.

Given that the declination error is heavy-tailed, as earlier published literature confirms, the conclusions based on the work in this thesis is that there is need for corrective actions regarding the existing methods. Whether or not these corrective actions are to change the statistical distributions on the error sources or they are minor changes in the existing methods will be a subject of further work.

8.2 Further work

If the work of this thesis should be a subject for further work, it must be done a more general analysis of the declination error. This thesis has focused on the effects of a heavy-tailed declination error, and not so much on finding a global declination error model. The chosen NIG-distribution for the declination error is based only on one processed annual serie of observations of the Earth's magnetic field on one location. The chosen annual serie has the characteristical properties of the declination error source which we seek, but there is need for further work on finding a more globally valid declination error model with focus on the heavy-tailed behaviours.

The errors in MWD magnetic measurements have been observed to be heavy-tailed. Based on published literature, we assumed that the declination error was the only error that had heavy tails. The declination error is in most cases the most significant error source, and it is also one of few error sources that it is possible to analyze directly from observations. If the error model should include other heavy tailed error sources, we need datasets of these errors in order to find a suitable NIG-distributions. Bias- and scale errors of accelerometers and magnetometers are possible to analyze by multi-station estimation. But there is no literature which indicate the need of heavy-tailed statistical distributions for these errors. Other error sources are hard, if not impossible to analyze, since there does not exist any data of these error sources.

The number of error sources that are included in our error models can easily be increased. An improved error model should include the use of correction algorithms for the axial magnetism in the drillstring. An apparent axial magnetism is corrected for by a new set of equations which connects the MWD-measurements with the measured depth, inclination and azimuth. This results in a new set of weighting functions which we, for implementation purposes, chose not to include in our model.

In our analyses we have assumed that the reference well and the offset well are uncorrelated. This is not the case in general, as a global error source influences both wells. We have seen that the well geometry is very important for the position uncertainty. Wells have often the same geometry in some sections as they can be drilled from the same template. In the case where the reference well is a sidestep of an offset well, we have a total correlation of both wells until the sidestep is made. Estimating the correlation between wells may be of interest for further work.

MWD magnetic directional surveying is a drilling operation which gives us information about the position while drilling. Between MWD-surveys or after drilling completion, it is possible to position the well using other surveying methods, such as with gyroscopic instruments. This gives us another error model of the position uncertainty than the one used in this thesis. In this thesis we have used the same error model on offset wells as on reference wells. Further work on this subject should also include the possibility to have other error models than the one we have used in this thesis.

Bibliography

[1] K. Aas. *Statistical modelling of financial risk.* PhD thesis, NTNU, 2008.

[2] K. Aas and I.H.Haff. Nig and skew student's t: Two special cases of the generalised hyperbolic distribution. *SAMBA*, 1, 2005.

[3] A. Azzalini and A. Capitanio. Statistical applications of the multivariate skew-normal distribution. *J. Roy. Statist. Soc.*, 61(3):579–602, 1998.

[4] A. Azzalini and A. Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *J. Roy. Statist. Soc.*, 65:367–389, 2003.

[5] A. G. Brooks and H. Wilson. An improved method for computing wellbore position uncertainty and its application to collision and target intersection probability analysis. *SPE European Petroleum Conference, Milan, Italy*, October 1996.

[6] B. T. Bruun. Effect of drillstring magnetism and sensor errors on mwd-directional surveys. Master's thesis, NTNU, 2000.

[7] G. Casella and R. L. Berger. *Statistical inference.* Duxbury advanced series. Duxbury, 2 edition, 2002.

[8] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *J. Roy. Statist. Soc.*, 39(1):1–38, 1977.

[9] R. Ekseth. *Uncertainties in connection with the determination of wellbore positions.* PhD thesis, NTNU, 1998.

[10] T. Gjerde. Application of heavy tailed distributions on measurement while drilling, 2007.

[11] T.A. Øigård, A. Hanssen, R.E. Hansen, and F. Godtliebsen. Em-estimation and modeling of heavy-tailed processes with the multivariate normal inverse gaussian distribution. *Signal Processing*, 85:1655–1673, 2005.

[12] D. Karlis. An em type algorithm for maximum likelihood estimation of the normal-inverse gaussian distribution. *Statistics and Probability letters*, 57:43–52, 2000.

[13] D. Karlis and A. Papadimitriou. Maximum likelihood estimation for the multivariate normal inverse gaussian model, September 2003.

[14] Landmark. *COMPASS for Windows Training Manual*.

[15] G. A. McNair, S. J. Lanc, J. Codling, and R. Watson. Implementation of a new risk based well collision avoidance method. *SPE/IADC Drilling conference, Amsterdam, Netherlands*, February 2005.

[16] R. T. Merrill. *The earth's magnetic field*, volume Volume 32 of *International geophysics series*. Academic press, 1983.

[17] E. Nyrnes. *Error analyses and quality control of wellbore directional surveys*. PhD thesis, NTNU, 2006.

[18] Tromsø Geophysical Observatory. http://geo.phys.uit.no/geomag.html.

[19] B.D. Ripley. *Stochastic simulation*. Wiley Series in probability and statistics, 1987.

[20] J. L. Thorogood and S.J. Sawaryn. A compendium of directional calculations based on the minimum curvature method. *SPE Drill and completion*, 20(1):24–36, 2005.

[21] T. Torkildsen and J. Bang. Directional surveying: Rotating and sliding operations give different wellbore position accuracy. *SPE Annual Technical Conference and Exhibition, Dallas, Texas*, October 2000.

[22] T. Torkildsen, R. H. Sveen, and J. Bang. Location dependent variations of declination. *IKU Petroleumsforskning AS*, 1997.

[23] T. Torkildsen, R. H. Sveen, and J. Bang. Time dependent variations of declination. *IKU Petroleumsforskning AS*, 1997.

[24] Wikipedia. http://en.wikipedia.org/wiki/.

[25] H. S. Williamson. Accuracy prediction for directional measurement while drilling. *SPE Drill and completion*, 15(4):221–233, 2000.

[26] H.S. Williamson. Towards risk-based well separation rules. *SPE Drill and completion*, 13(1):47–51, 1998.

Appendix A

Proof: Linearity formulas of the normal inverse gaussian distribution

Let $X \sim NIG(\mu, \rho, \alpha, \beta)$. The moment-generating function of X is:

$$M_X(t) = \exp\left\{\mu t + \rho(\gamma - \sqrt{\alpha^2 + (\beta + t)^2})\right\} \tag{A.1}$$

Let $Y = aX + b$. From the properties of the moment-generating functions, we can calculate the moment-generating function of Y:

$$\begin{aligned}
M_Y(t) &= e^{bt} M_X(at) \\
&= \exp\left\{bt + \mu at + \rho(\gamma - \sqrt{\alpha^2 - (\beta + at)^2})\right\} \\
&= \exp\left\{(a\mu + b)t + \rho(\gamma - \sqrt{\alpha^2 - (\beta + at)^2})\right\} \\
&= \exp\left\{(a\mu + b)t + |a|\rho(\frac{\gamma}{|a|} - \frac{1}{|a|}\sqrt{\alpha^2 - (\beta + at)^2})\right\} \\
&= \exp\left\{(a\mu + b)t + |a|\rho(\frac{\gamma}{|a|} - \sqrt{\frac{\alpha^2}{a^2} - \frac{1}{a^2}(\beta + at)^2})\right\} \\
&= \exp\left\{(a\mu + b)t + |a|\rho(\frac{\gamma}{|a|} - \sqrt{\frac{\alpha^2}{a^2} - (\frac{\beta}{a} + t)^2})\right\} \\
&= \exp\left\{(a\mu + b)t + |a|\rho(\frac{\gamma}{|a|} - \sqrt{\left(\frac{\alpha}{|a|}\right)^2 - (\frac{\beta}{a} + t)^2})\right\}
\end{aligned}$$

We see that the moment-generating function of Y is of the same form as (A.1), i.e. $Y \sim NIG(a\mu + b, |a|\rho, \frac{\alpha}{|a|}, \frac{\beta}{a})$.

Let now $Y = \sum_{i=1}^{n} X_i$, where $X_i \sim NIG(\mu_i, \rho_i, \alpha, \beta)$. The moment-generating function of Y can be expressed by the moment-generating functions of X_i:

$$\begin{aligned}
M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) \\
&= \exp\left\{ \sum_{i=1}^{n} \Big(\mu_i t + \rho_i [\gamma - \sqrt{\alpha^2 - (\beta + t)^2}] \Big) \right\} \\
&= \exp\left\{ \left(\sum_{i=1}^{n} \mu_i \right) t + \left(\sum_{i=1}^{n} \rho_i \right) [\gamma - \sqrt{\alpha^2 - (\beta + t)^2}] \right\}
\end{aligned}$$

We see that $Y \sim NIG(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \rho_i, \alpha, \beta)$.

Appendix B

Proof: The normal inverse gaussian distribution as a normal variance mean-mixture

Let Z be a Inverse Gaussian (IG) distributed random variable, $Z \sim IG(\rho, \gamma)$. The density function of Z is:

$$f_Z(z) = \frac{1}{\sqrt{2\pi z^3}} \rho \exp\left\{\rho\gamma - \frac{1}{2}(\frac{\rho^2}{z} + \gamma^2 z)\right\}$$

Let $X|Z = z \sim N(\mu + \beta z, z)$. The joint density of (X, Z) is:

$$\begin{aligned}
f(x,z) &= f(x|z) \times f(z) \\
&= \frac{1}{\sqrt{2\pi z}} \exp\left\{-\frac{1}{2z}(x - (\mu + \beta z))^2\right\} \times \frac{1}{\sqrt{2\pi z^3}} \rho \exp\left\{\rho\gamma - \frac{1}{2}(\frac{\rho^2}{z} + \gamma^2 z)\right\} \\
&= \frac{1}{2\pi z^2} \rho \exp\left\{\rho\gamma - \frac{1}{2z}(x - (\mu + \beta z))^2 - \frac{1}{2}(\frac{\rho^2}{z} + \gamma^2 z)\right\} \qquad \text{(B.1)} \\
&\propto \rho z^{-2} e^{\rho\gamma - \beta\mu} \exp\left\{\beta x + \mu\frac{x}{z} - \frac{1}{2}(\beta^2 + \gamma^2) z - \frac{1}{2}(\mu^2 + \rho^2) z^{-1}\right\} \qquad \text{(B.2)}
\end{aligned}$$

We compute the marginal density of X from the joint density (B.1):

$$
\begin{aligned}
f_X(x) &= \int_0^\infty f(x|z)f(z)\mathrm{d}z \\
&= \int_0^\infty \frac{1}{2\pi z^2}\rho\exp\left\{\rho\gamma - \frac{1}{2z}(x-(\mu+\beta z))^2 - \frac{1}{2}(\frac{\rho^2}{z}+\gamma^2 z)\right\}\mathrm{d}z \\
&= \frac{1}{\pi}\rho\exp\left\{\rho\gamma+\beta(x-\mu)\right\}\times \\
&\quad\times\frac{1}{2}\int_0^\infty \frac{1}{z^2}\exp\left\{-\frac{1}{2}[\frac{\rho^2}{z}+\gamma^2 z+\frac{1}{z}(x^2-2\mu x+\mu^2+\beta^2 z^2)]\right\}\mathrm{d}z \\
&= \frac{1}{\pi}\rho\exp\left\{\rho\gamma+\beta(x-\mu)\right\}\frac{1}{2}\int_0^\infty \frac{1}{z^2}\exp\left\{-\frac{1}{2}[\frac{\rho^2}{z}+\alpha^2 z+\frac{1}{z}(x-\mu)^2]\right\}\mathrm{d}z \\
&= \frac{1}{\pi}\rho\exp\left\{\rho\gamma+\beta(x-\mu)\right\}\frac{1}{2}\int_0^\infty \frac{1}{z^2}\exp\left\{-\frac{1}{2}[\frac{1}{z}(\rho^2+(x-\mu)^2)+\alpha^2 z]\right\}\mathrm{d}z
\end{aligned}
$$

We substitute $u=\frac{\sqrt{\rho^2+(x-\mu)^2}}{\alpha z}$:

$$
\begin{aligned}
f_X(x) &= \frac{1}{\pi}\rho\exp\left\{\rho\gamma+\beta(x-\mu)\right\}\times \\
&\quad\times\frac{1}{2}\int_0^\infty \exp\left\{-\frac{1}{2}\alpha\sqrt{\rho^2+(x-\mu)^2}[u+\frac{1}{u}]\right\}\frac{\alpha}{\sqrt{\rho^2+(x-\mu)^2}}\mathrm{d}u \\
&= \frac{\alpha\rho}{\pi\sqrt{\rho^2+(x-\mu)^2}}\;\exp\left\{\rho\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)\right\}K_1\left(\alpha\sqrt{\rho^2+(x-\mu)^2}\right)
\end{aligned}
$$

The last expression is in fact the NIG density function (3.12) with parameters (μ,ρ,α,β), as we recognize that the last integral is the modified Bessel function of order 1.

Appendix C

Normal approximation in anti-collision hypothesis testing

The hypothesis test in Section 6.4.1 approximate the distribution of D with a normal distribution. The standard deviation of D, σ_D, is approximated as shown in (6.2). σ_D is dependent on the covariance structure in both wells, which makes it hard to prove in general if the normal approximation is good. We can use simulations to check the normal approximation. From Section 6.4 we have the following difference vector of interest:

$$r \sim N_3(u_0 - v_0, \Sigma_u + \Sigma_v)$$

For different u_0, v_0, Σ_u and Σ_v, we can simulate realizations of r, and compute realizations of the distance $D = \sqrt{r^T r}$. We show three different simulations in Figure C.1. The conclusion of these simulations is that the normal approximation is good when the true μ_D is large relative to the standard deviation σ_D, as shown in the middle simulation in Figure C.1. In the case where we have a relatively high risk of well-collision, the left tail of the fitted normal distribution is defined for negative values of D, which is the case for the upper and lower simulation in Figure C.1. The probability of well-collision is usually small in practice, which would result in a simulation with few small values of D. So the normal approximation is valid when it is not a high risk of well-collision.

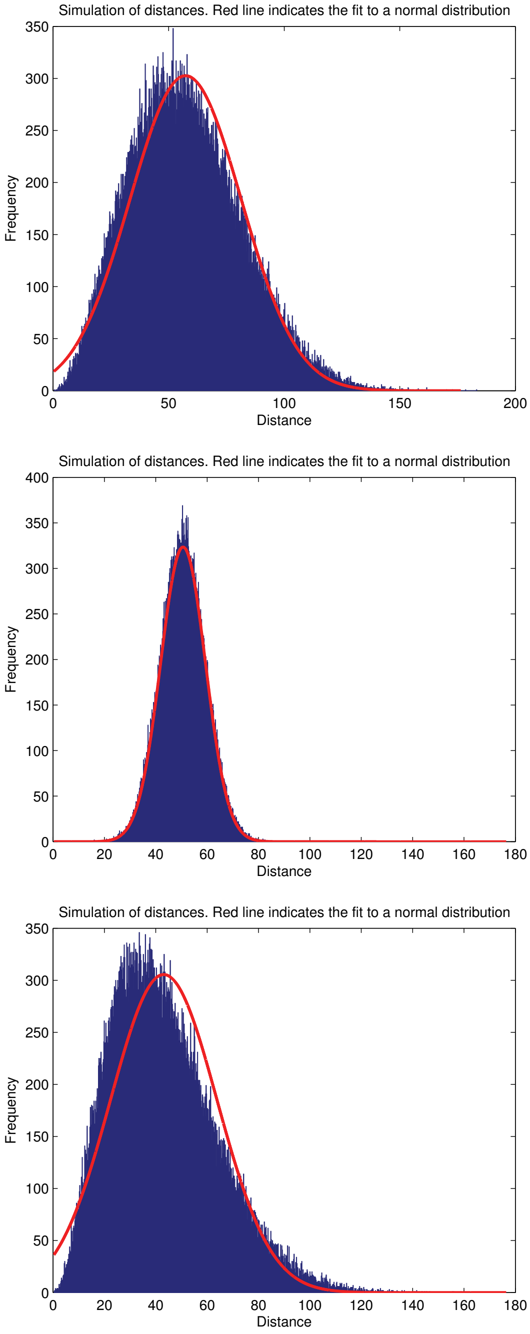


Figure C.1: Three different simulations of the distance D. The upper simulation has large expected value and large variance. The middle simulation has large expected value and small variance. The lower simulation has small expected value and large variance. The normal distribution is a good approximation when D is not too small relative to the uncertainty in D.

Appendix D

An alternative hypothesis test in anti-collision calculations

An alternative hypothesis test to the tests in Section 6.4.1 and 6.4.2 can be used in anti-collision calculations. Under the assumption that the position uncertainties are multivariate normally distributed, $u \sim N_3(u_0, \Sigma_u)$ and $v \sim N_3(v_0, \Sigma_v)$, we also know the distribution of the difference vector r. $r \sim N_3(r_0, \Sigma_r)$, where $r_0 = u_0 - v_0$ and $\Sigma_r = \Sigma_u + \Sigma_v$. We can use the following hypothesis test when we neglect the wellbore diameters:

$$\begin{aligned} H_0 \quad &: \quad r_0 = 0 \\ H_1 \quad &: \quad r_0 \neq 0 \end{aligned}$$

In this hypothesis test, 0 means the null vector. Under H_0 we have that $r \sim N(0, \Sigma_r)$. We can then use $r^T(\Sigma_r)^{-1}r \sim \chi_3^2$ as our test statistic. We reject H_0 on significance level α if $r_0^T(\Sigma_r)^{-1}r_0 > k'_\alpha$, where k'_α is the $100(1-\alpha)$-quantile of the χ_3^2-distribution. This hypothesis test is not easily applicable when we introduce the wellbore diameters. Then we would have to do a test on the length of r_0, as follows:

$$\begin{aligned} H_0 \quad &: \quad ||r_0|| \leq \frac{d_u + d_v}{2} \\ H_1 \quad &: \quad ||r_0|| > \frac{d_u + d_v}{2} \end{aligned}$$

This hypothesis test is not as easily practiced as the hypothesis tests in Section 6.4.1 and 6.4.2, since we can not say anything about the expected vector of r under H_0, as we could in the special case of $||r_0|| = 0$. This test may be used where there are large position uncertainties relative to the wellbore diameters.