The GNSS integer ambiguities: estimation and validation

Sandra Verhagen

The GNSS integer ambiguities: estimation and validation

Sandra Verhagen

NCG Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission

Delft, January 2005

The GNSS integer ambiguities: estimation and validation
Sandra Verhagen
Publications on Geodesy 58
ISBN 90 6132 290 1
ISSN 0165 1706

Published by: NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission, Delft, The Netherlands
Printed by: Optima Grafische Communicatie, Optima Graphic Communication, Rotterdam, The Netherlands
Cover illustration: René Buur

NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission
P.O. Box 5058, 2600 GB Delft, The Netherlands
T: +31 (0)15 278 28 19
F: +31 (0)15 278 17 75
E: ncg@lr.tudelft.nl
W: www.ncg.knaw.nl

The NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission is part of the Royal Netherlands Academy of Arts and Sciences (KNAW)

Precision is a quest on which travelers journey halfway to their destination, and then halfway again and again and again, never reaching finality.
ZENO

from: Ken Alder, *The measure of all things*

Contents

Preface

It was in the beginning of the nineties that the first major steps for solving the problem of GPS integer ambiguity resolution were made. Ten years later part of the problem could be considered solved, as efficient algorithms were available for the estimation of the integer ambiguities with the highest possible success rate. However, as a geodesist one cannot be content with this result, since parameter resolution means estimation *and* validation. The latter problem turned out to be very complex due to the non-standard distributions of the parameters involved.

Prof. Peter Teunissen has made a major contribution to the solution of the integer estimation problem with the introduction of the LAMBDA method in 1993. He made me aware of the open problem with respect to integer validation, and - as I started to work on this problem - has given me indispensable theoretical input.

This thesis starts with a description of the GNSS observation models, and then outlines the theory on integer estimation and validation as it was available when I started to work on this topic. This part does contain some relatively new results on the parameter distributions, and some important comments on the validation procedures as proposed in the past and used in practice. The remaining chapters are all devoted to 'new' estimation and validation algorithms.

In this place I would of course like to thank all of my (former) colleagues at MGP for their support and contribution to a good research environment. A special thanks goes to the following people: my promotor Peter Teunissen for the supervision, input, feedback, comments and discussions; Kees de Jong, my supervisor in the first two years; Dennis Odijk for proofreading this thesis; Peter Joosten for the computer support; Ria Scholtes for the administrative support; the members of the examination committee for their critical reviews and useful comments.

Finally, I would like to thank René for his help with the layout and graphics and of course his overall support.

Summary

The GNSS integer ambiguities: estimation and validation

Fast and high precision relative positioning with a Global Navigations Satellite System (GNSS) is only possible by using the very precise carrier phase measurements. However, these carrier phases are ambiguous by an unknown number of cycles. The knowledge that the ambiguities are integer-valued has been exploited in the past 15 years for the development of integer ambiguity resolution algorithms. Once the ambiguities are fixed to their integer values, the carrier phase measurements start to act as if they were very precise pseudorange measurements.

The estimation process consists then of three steps. First a standard least-squares adjustment is applied in order to arrive at the so-called float solution. All unknown parameters are estimated as real-valued. In the second step, the integer constraint on the ambiguities is considered. This means that the float ambiguities are mapped to integer values. Different choices of the map are possible. The float ambiguities can simply be rounded to the nearest integer values, or conditionally rounded so that the correlation between the ambiguities is taken into account. The optimal choice is to use the integer least-squares estimator, which maximizes the probability of correct integer estimation. Finally, after fixing the ambiguities to their integer values, the remaining unknown parameters are adjusted by virtue of their correlation with the ambiguities.

Nowadays, the non-trivial problem of integer ambiguity estimation can be considered solved. However, a parameter estimation theory is not complete without the appropriate measures to validate the solution. So, fixing the ambiguities should only be applied if there is enough confidence in their correctness. The probability of correct integer estimation can be computed a priori – without the need for actual observations – and is called the success rate. Only if this success rate is very close to one, the estimated fixed ambiguities may be considered deterministic. In that case it is possible to define test statistics in order to validate the fixed solution. If the success rate is not close to one, in practice a user has to choose between two undesirable situations:

- Integer validation is based on wrong assumptions, since the randomness of the fixed ambiguities is incorrectly ignored;

- No attempt is made to fix the ambiguities because of the low success rate, although there is still a probability that the ambiguities can be fixed correctly.

Examples of integer validation tests proposed in literature, are the ratio test, the projector test, and the difference test. Some of these integer validation tests are defined such that the invalid assumption of deterministic fixed ambiguities is avoided. However, these tests lack a sound theoretical foundation. Moreover, often fixed critical values are used based on experience. But that implies that in many situations the tests will either be too conservative or too optimistic.

Obviously, validation of the integer ambiguity solution is still an open problem. In order to avoid the above two situations, two approaches are investigated here.

The first method uses a new ambiguity estimator, which is in some sense always superior to its float and fixed counterparts. This estimator is the *Best Integer Equivariant* (BIE) estimator. It is best in the sense that it minimizes the mean squared errors of the estimators. This is a weaker performance criterion than that of the integer least-squares estimator, the maximization of the success rate. This might be an advantage if the success rate is not high, since then a user would not use the fixed solution, but stick to the float solution. The BIE ambiguity estimator is equal to a weighted sum over all integer vectors, and the weights depend on the probability density function of the float ambiguities. Therefore, the BIE estimator can be considered as a compromise between the float and fixed solution; those are approximated in the extreme cases of very bad or very good precisions, respectively. The disadvantages of BIE estimation are that an exact probabilistic evaluation of the solution is still not possible, and the procedure is computationally complex. Moreover, it was shown that the BIE estimator significantly outperforms the float or fixed solution only in a limited number of cases.

Therefore, another new class of integer estimators is investigated: the class of *Integer Aperture* (IA) estimators. IA estimation is an overall approach of integer estimation and validation. This is the case because three a priori probabilities are distinguished: the success rate, the fail rate, and the undecided rate. Around each integer an identically shaped acceptance region, referred to as the aperture pull-in region, is defined, with the size determined by the choice of a maximum allowed fail rate. With respect to the shape of the regions different choices are possible. Examples are ellipsoids and scaled versions of the integer bootstrapping or integer least-squares pull-in regions. It is also shown that the acceptance regions corresponding to the ratio tests, the projector test and the difference test are examples of aperture pull-in regions. Finally, the *optimal* IA estimator is defined such that the success rate is maximized for a fixed fail rate.

IA estimation has several important advantages. The first is that it is an overall approach to the integer estimation and validation problem, which makes exact probabilistic evaluation of the final solution possible. Furthermore, the acceptance region is determined by the choice of the maximum allowed fail rate, which is exactly what a user wants. This means that the region depends on the model at hand, and thus on the precision. With the traditional approaches this is not the case, so that in general the tests are either too conservative or too optimistic. In the first case this leads to a unnecessarily long time to first fix, in the second case the fail rate will be unacceptably high.

Finally, it is shown that the popular ratio test and the difference test perform almost as good as the optimal IA estimator if the fixed fail rate approach is used. So, a theoretical justification is given for these tests, such that an appropriate choice of the critical values is possible. Also the IA least-squares estimator, based on using down-scaled integer least-squares pull-in regions, performs well. This estimator is identical to the integer least-squares estimator combined with integer hypothesis testing, but with the advantage that the acceptance region is determined by the choice of the fail rate and not by the false alarm rate.

Samenvatting (in Dutch)

De GNSS geheeltallige meerduidigheden: schatting en validatie

Hoge precisie relatieve plaatsbepaling met behulp van een Global Navigation Satellite System (GNSS) is alleen mogelijk als de zeer precieze fasewaarnemingen van de draaggolf gebruikt worden. Helaas zijn deze fasewaarnemingen meerduidig, omdat niet gemeten kan worden hoeveel cycli er aan de waarneming zijn voorafgegaan. Wel is bekend dat de meerduidigheden geheeltallig zijn, en deze kennis is uitgebuit voor de ontwikkeling van algoritmes voor geheeltallige meerduidigheidsbepaling. Als de meerduidigheden eenmaal geschat zijn, kunnen de fasewaarnemingen beschouwd worden als zeer precieze pseudo-afstandsmetingen.

De schattingsprocedure bestaat dan uit drie stappen. Allereerst wordt er een standaard kleinstekwadratenvereffening uitgevoerd, resulterend in de zogenaamde *float* oplossing. Alle onbekende parameters worden in dit geval geschat als reëelwaardig. In de tweede stap wordt er gebruik gemaakt van de voorwaarde dat de meerduidigheden geheeltallig moeten zijn. Daarvoor kunnen verschillende schatters gebruikt worden. Het simpelst is om de float meerduidigheden af te ronden naar de dichtstbijzijnde geheeltallige waarden (*integer rounding*). Een andere mogelijkheid is om de meerduidigheden conditioneel af te ronden, zodat de correlatie van de meerduidigheden betracht wordt (*integer bootstrapping*). De optimale keuze is echter om de *integer* kleinstekwadraten schatter te gebruiken, omdat deze de kans op correcte geheeltallige schatting maximaliseert. Tenslotte, nadat de meerduidigheden bepaald zijn, worden de resterende parameters gecorrigeerd op basis van hun correlatie met de meerduidigheden. De uiteindelijke oplossing wordt de *fixed* oplossing genoemd.

Het niet-triviale probleem van geheeltallige meerduidigheidsschatting kan op dit moment als opgelost beschouwd worden. Echter, zonder geschikte maten om de schatting te valideren, is de oplossing nog niet compleet. De fixed meerduidigheidsoplossing mag alleen gebruikt worden als er genoeg vertrouwen is in de correctheid ervan. De kans op het correct schatten van de geheeltallige meerduidigheden kan a priori, dat wil zeggen zonder dat waarnemingen nodig zijn, berekend worden. Deze kans wordt de *success rate* genoemd. Alleen als de success rate zeer dichtbij één ligt, mogen de geschatte fixed meerduidigheden als deterministisch beschouwd worden. In dat geval is het mogelijk om toetsgrootheden te definiëren die validatie van de fixed oplossing mogelijk maken. Maar als de success rate te laag is, zal een gebruiker in de praktijk een keuze moeten maken tussen twee onwenselijke situaties:

- Validatie van de geheeltallige oplossing gebaseerd op foute aannames, doordat de fixed meerduidigheden onterecht als deterministisch beschouwd worden;

- Er wordt geen poging gedaan de meerduidigheden te bepalen omdat de success rate te laag is, terwijl er toch een kans bestaat dat de juiste geheeltallige oplossing gevonden zou worden.

Voorbeelden van integer validatietoetsen die in de literatuur zijn voorgesteld zijn de ratio-toets, de projectortoets en de verschiltoets. Sommige van deze integer validatie toetsen zijn zodanig gedefinieerd dat de foutieve aanname van deterministische fixed meerduidigheden wordt vermeden. Deze toetsen ontberen echter een goede theoretische basis. Bovendien worden vaak vaste kritieke waarden gebruikt die gebaseerd zijn op ervaring. Maar dat betekent dat in veel situaties de toetsen ofwel te conservatief ofwel te optimistisch zullen zijn.

In dit proefschrift is het onderzoek beschreven naar twee methodes om het geschetste probleem op te lossen.

De eerste methode betreft het gebruik van een nieuwe schatter, de zogenaamde *best integer equivariant* (BIE) schatter. Deze schatter maakt wel gebruik van de voorwaarde van geheeltallige meerduidigheden, maar niet op de optimale manier. De schatter is 'best' in de zin dat de mean-squared error (MSE) geminimaliseerd wordt. De BIE meerduidigheidsschatter is gelijk aan een gewogen som over alle integer vectoren, waarbij de gewichten afhankelijk zijn van de kansdichtheidsfunctie van de float meerduidigheden. De BIE schatter kan daarom beschouwd worden als een compromis tussen de float en fixed oplossing; deze worden benaderd in de extreme gevallen van respectievelijk oneindig slechte of oneindig goede precisie. Nadelen van BIE schatting zijn dat een exacte probabilistische evaluatie van de oplossing nog steeds niet mogelijk is, en dat de procedure nogal complex is en veel rekentijd vergt. Bovendien is aangetoond dat de BIE schatter slechts in een beperkt aantal gevallen tot duidelijk betere resultaten leidt dan de float of de fixed oplossing.

Daarom is ook onderzoek gedaan naar een andere nieuwe klasse van schatters: de klasse van *integer aperture* (IA) schatters. Hierbij zijn schatting en validatie geïntegreerd, doordat a priori drie kansen worden onderscheiden: de kans op correct fixen (success rate), de kans op foutief fixen (fail rate), en de kans op niet fixen (undecided rate). Rondom elke integer wordt een gelijkvormig acceptatiegebied (aperture pull-in region) gedefinieerd, waarbij de grootte bepaald wordt door de keuze van een maximaal toelaatbare fail rate. Met betrekking tot de vorm zijn verschillende keuzes mogelijk. Voorbeelden zijn ellipsoïdes en geschaalde pull-in regions (bootstrapping of integer least-squares). Verder is aangetoond dat ook de acceptatiegebieden behorend bij de ratio-toetsen, de projectortoets en de verschiltoets voorbeelden zijn van aperture pull-in regions. Tot slot is de optimale IA schatter gedefinieerd zodanig dat voor een vaste fail rate de success rate wordt gemaximaliseerd.

IA schatting heeft belangrijke voordelen. Ten eerste is het een *overall* aanpak voor het integer schattings- en validatieprobleem, die tevens een exacte probabilistische evaluatie van de uiteindelijke oplossing mogelijk maakt. Verder wordt het acceptatiegebied vast-

gelegd door de keuze van de maximaal toelaatbare fail rate, precies wat een gebruiker wil. Dit betekent dat het acceptatiegebied afhangt van het model, en dus de precisie, waarmee op dat moment gewerkt wordt. Met de traditionele methodes is dat meestal niet het geval en zijn de toetsen ofwel te conservatief of te optimistisch. In het eerste geval leidt dat tot onnodig lange 'times to first fix', in het tweede geval zal de fail rate onacceptabel hoog zijn.

Tot slot is aangetoond dat de populaire ratio toets en de verschiltoets bijna net zo goed presteren als de optimale IA schatter indien de kritieke waarde wordt bepaald door de maximaal toelaatbare fail rate. Er is nu dus een theoretische basis beschikbaar voor het gebruik van deze toetsen, zodanig dat ook een correcte keuze van de kritieke waarde mogelijk is. Ook de IA kleinstekwadraten schatter, gebaseerd op het gebruik van geschaalde integer least-squares pull-in regions, presteert goed. Deze schatter is identiek aan de integer least-squares schatter in combinatie met de integer hypothese toetsing zoals hiervoor beschreven, maar met als voordeel dat het acceptatiegebied bepaald wordt door de keuze van de fail rate en niet van de kans op vals alarm.

Notation and symbols

Mathematical notation and matrices

I_p	identity matrix of order p
e_p	p-vector with ones
0	vector or matrix with zeros
$\|Q\|$	determinant of matrix Q
$\|x\|_Q^2$	squared norm, with positive definite Q: $\|x\|_Q^2 = x^T Q^{-1} x$
$[x]$	x is rounded to the nearest integer
C_z^ϵ	ellipsoidal region centered at z and with radius ϵ

Signals

c	velocity of light in vacuum: $c = 299792458$ m s^{-1}
f	frequency
λ	wavelength: $\lambda = c/f$

GNSS observation model

$\square^s$	refers to satellite s
$\square_r$	refers to receiver r
$\square_{,j}$	refers to frequency j
m	number of satellites
f	number of frequencies
k	number of epochs
p, ϕ	code and phase observation respectively
t^s	transmission time from satellite s
t_r	reception time at receiver r
τ	signal travel time
e, ε	code and phase noise respectively
$dt, \delta t$	clock error of code and phase observations respectively

d, δ	instrumental delay for code and phase observations respectively
$da, \delta a$	atmospheric delay for code and phase observations respectively
T	tropospheric delay
I	ionospheric delay
μ_j	$= \lambda_j^2 / \lambda_1^2$
$dm, \delta m$	multipath effect for code and phase observations respectively
M	real-valued carrier phase ambiguity parameter
N	integer-valued carrier phase ambiguity parameter
ρ	satellite-receiver range
$\mathbf{r}$	position vector
$\mathbf{u}$	unit line-of-sight vector
ψ	mapping function for tropospheric delays
P, Φ	vector with code and phase observation respectively
ρ	vector with satellite-receiver ranges
T	vector with tropospheric delays
$\mathcal{I}$	vector with ionospheric delays
$\mathbf{dt}$	vector with lumped instrumental delays and clock errors
a	vector with carrier phase ambiguities
Λ	diagonal matrix with wavelengths as diagonal elements
G	matrix with line-of-sight vectors
Ψ	vector with mapping functions for tropospheric delays
D	double difference operator
σ_p, σ_ϕ	standard deviation of code and phase observation respectively
σ_I	standard deviation of ionospheric pseudo-observation

Estimation

$E\{\cdot\}$	expectation operator
$D\{\cdot\}$	dispersion operator
y	m-vector of observations
a	n-vector with ambiguity parameters
b	p-vector with real-valued parameters: baseline, atmospheric delays, etc.
e, ε	noise parameters
A, B	design matrices: $E\{y\} = Aa + Bb + e$
Q_x	variance-covariance matrix of x
$Q_{x\|v}$	conditional variance-covariance matrix: $Q_{x\|v} = Q_x - Q_{xv} Q_v^{-1} Q_{vx}$
σ^2	variance factor
G_x	cofactor matrix of x: $Q_x = \sigma^2 G_x$
$Q = LDL^T$	LDL^T-decomposition of Q with lower triangular matrix L, diagonal matrix D
$\sigma_{i\|I}$	conditional standard deviation of ith entry of a parameter vector conditioned on the $I = 1, \ldots, i-1$ entries

Parameter distributions

$f_{\hat{x}}(x)$	probability density function of random variable $\hat{x}$
$f_{\hat{x}\|\hat{v}}(x\|v)$	conditional probability density function of random variable $\hat{x}$
$\hat{x} \sim N(x, Q_x)$	random variable $\hat{x}$ is normally distributed with mean x and variance-covariance matrix Q_x
$\hat{x} \sim \chi^2(n, \lambda)$	random variable $\hat{x}$ has χ^2-distribution with n degrees of freedom and non-centrality parameter λ
$\hat{x} \sim F(m, n, \lambda)$	random variable $\hat{x}$ has F-distribution with m and n degrees of freedom and non-centrality parameter λ
$P(\hat{x} = x)$	probability that $\hat{x}$ will be equal to the mean x
$P(\hat{x} = x\|\hat{v} = v)$	probability that $\hat{x}$ will be equal to the mean x conditioned on $\hat{v} = v$
α	false alarm rate or level of significance
γ	detection power

Ambiguity resolution

$\hat{x}$	float estimate of x (= least-squares estimate)
$\check{x}$	fixed estimate of x
$\check{\epsilon}$	ambiguity residuals $\hat{a} - \check{a}$
$\tilde{x}$	best integer equivariant estimate of x
$\bar{x}$	integer aperture estimate of x
P_s, P_f, P_u	probability of success, failure, and undecided respectively
$P_{s,\mathrm{LS}}, P_{f,\mathrm{LS}}$	integer least-squares success and fail rate respectively
S_z	pull-in region centered at the integer z
$s_z(x)$	indicator function: $s_z(x) = 1 \Leftrightarrow x \in S_z$, $s_z(x) = 0$ otherwise
$\hat{\Omega}$	squared norm of residuals of float solution: $\hat{\Omega} = \hat{e}^T G_y^{-1} \hat{e}$
$\hat{\sigma}^2$	float estimate of variance factor: $\hat{\sigma}^2 = \frac{\hat{\Omega}}{m-n-p}$
$\check{\Omega}$	squared norm of residuals of fixed solution: $\check{\Omega} = \check{e}^T G_y^{-1} \check{e}$
$\check{\sigma}^2$	fixed estimate of variance factor: $\check{\sigma}^2 = \frac{\check{\Omega}}{m-p}$
R_i	squared norm of ambiguity residuals: $R_1 = (\hat{a} - \check{a})^T G_{\hat{a}}^{-1} (\hat{a} - \check{a}) = \check{\Omega} - \hat{\Omega}$ $R_2 = (\hat{a} - \check{a}_2)^T G_{\hat{a}}^{-1} (\hat{a} - \check{a}_2)$
Ω_z	aperture pull-in region centered at the integer z
Ω	aperture space: $\Omega = \bigcup_{z \in \mathbb{Z}^n} \Omega_z$
μ	aperture parameter

Acronyms

ADOP	Ambiguity Dilution Of Precision
B	integer Bootstrapping
BIE	Best Integer Equivariant
CIR	Cascade Integer Resolutoin
DIA	Detection-Identification-Adaptation
DD	Double Difference
DGNSS	Differential GNSS
DGPS	Differential GPS
EIA	Ellipsoidal Integer Aperture
GLONASS	GLObal NAvigation Satellite System (Russian)
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IA	Integer Aperture
IAB	Integer Aperture Bootstrapping
IALS	Integer Aperture Least-Squares
IE	Integer Equivariant
ILS	Integer Least-Squares
ITCAR	Integrated TCAR
LAMBDA	Least-squares AMBiguity Decorrelation Adjustment
LOM	Local Overall Model (test)
LOS	Line-Of-Sight
LS	Least-Squares
MDB	Minimal Detectable Bias
MDE	Minimal Detectable Effect
OIA	Optimal Integer Aperture
OTF	On-The-Fly
PDF	Probability Density Function
PIA	Penalized Integer Aperture
PMF	Probability Mass Function
R	integer Rounding
RTK	Real-Time Kinematic
SD	Single Difference
TCAR	Three Carrier Ambiguity Resolution
TEC	Total Electron Content
ZTD	Zenith Tropospheric Delay

1 Introduction

1.1 Background

In the past decade the Global Positioning System (GPS) has found widespread use in different fields such as surveying, navigation and geophysics. The requirements on precision, reliability and availability of the navigation system for these applications have become higher and higher. Moreover, the positioning information often needs to be obtained in (near) real-time. Without the development of integer ambiguity resolution algorithms, it would never have been possible to meet these requirements with GPS.

Fast and high precision relative positioning with a Global Navigations Satellite System (GNSS) is namely only possible by using the very precise carrier phase measurements. However, these carrier phases are ambiguous by an unknown number of cycles. The ambiguities are known to be integer-valued, and this knowledge has been exploited for the development of integer ambiguity resolution algorithms. Once the ambiguities are fixed on their integer values, the carrier phase measurements start to act as if they were very precise pseudorange measurements.

The estimation process consists then of three steps. First a standard least-squares adjustment is applied in order to arrive at the so-called float solution. All unknown parameters are estimated as real-valued. In the second step, the integer constraint on the ambiguities is considered. This means that the float ambiguities are mapped to integer values. Different choices of the map are possible. The float ambiguities can simply be rounded to the nearest integer values, or conditionally rounded so that the correlation between the ambiguities is taken into account. The optimal choice is to use the integer least-squares estimator, which maximizes the probability of correct integer estimation. Finally, after fixing the ambiguities to their integer values, the remaining unknown parameters are adjusted by virtue of their correlation with the ambiguities.

Nowadays, the non-trivial problem of integer ambiguity estimation can be considered solved. However, a parameter estimation theory is not complete without the appropriate measures to validate the solution. So, fixing the ambiguities should only be applied if there is enough confidence in their correctness. The probability of correct integer estimation can be computed a priori – without the need for actual observations – and is called the success rate. Only if this success rate is very close to one, the fixed ambiguity estimator may be considered deterministic. In that case it is possible to define test statistics in order to validate the fixed solution. If the success rate is not close to one,

in practice, a user has to choose between two undesirable situations:

- Integer validation is based on wrong assumptions, since the randomness of the fixed ambiguities is incorrectly ignored;
- No attempt is made to fix the ambiguities because of the low success rate, although there is still a probability that the ambiguities can be fixed correctly.

Some integer validations tests are defined such that the invalid assumption of a deterministic fixed ambiguity estimator is avoided. However, these tests lack a sound theoretical foundation. Moreover, often fixed critical values are used based on experience. But that implies that in many situations the tests will either be too conservative or too optimistic.

1.2 Objectives and contribution of this work

Obviously, validation of the integer ambiguity solution is still an open problem. In order to deal with the above two situations, two approaches are investigated:

- *A new class of ambiguity estimators is used, which results in estimators that in some sense are always superior to their float and fixed counterparts.*
  This estimator is the *Best Integer Equivariant* (BIE) estimator. It is best in the sense that it minimizes the mean squared errors of the estimators. This can be considered a weaker performance criterion than that of the integer least-squares estimator, the maximization of the success rate. This might be an advantage if the success rate is not high, since then a user will generally not have enough confidence in the fixed solution, and stick to the float solution. It is investigated how BIE estimation can be implemented. The performance of this estimator is then compared with the fixed and float estimators.
- *A new ambiguity estimator is used by defining an a priori acceptance region, or aperture space.*
  This new ambiguity resolution method is the so-called *integer aperture estimation*. The problem of ambiguity validation is incorporated in the estimation procedure with this approach. Instead of distinguishing only the probability of success and of failure, also the probability of not fixing (undecided rate) is considered a priori. An integer acceptance region is then defined by putting constraints on the three probabilities; only if the float ambiguities fall in the acceptance region, the ambiguities will be fixed using one of the well-known integer estimators.

The theoretical aspects of all approaches are considered, as well as implementation aspects. Thereby it is avoided to go into too much detail with respect to proofs and derivations; these are only included if they contribute to the understanding of the theory. The methods are implemented in Matlab®.

Furthermore, it is investigated how the various methods will perform. For that purpose, Monte-Carlo simulations are used, since then the 'true' situation is known and it is possible to compare the performance of different estimators and validators.

The contribution of this thesis is that a clear overview of the traditional and the new ambiguity resolution methods will be given, and it will be investigated why and when a certain approach is to be preferred – from a theoretical point of view as well as by comparing the performance of the methods. Furthermore, the implementation aspects will be discussed, as well as the open issues in that respect.

1.3 Outline

The outline of this thesis is as follows. In chapter 2 a brief introduction to GNSS positioning is given. The observations models are presented, as well as the general theory of quality control for GNSS applications.

Chapter 3 is devoted to the problem of ambiguity resolution. Three integer estimators are described: integer rounding, integer bootstrapping, and integer least-squares. Furthermore, the probabilistic properties of the float and fixed estimators are analyzed. Finally, the existing methods for integer validation are described and evaluated. The pitfalls of these methods, both theoretical and practical, will be discussed.

Integer least-squares estimation is generally accepted as the method to be used for integer ambiguity estimation. A disadvantage is, however, that it is not guaranteed that the resulting baseline estimator is closer to the true but unknown baseline than the original least-squares (float) estimator. Therefore, it is investigated in chapter 4 how the *best integer equivariant* estimator performs, which is defined such that the resulting baseline estimator always outperforms the least-squares solution in a certain sense.

The conclusions on the existing integer validation methods reveal the need for an integer testing method, which is based on a sound theoretical criterion. For that purpose, an alternative ambiguity resolution method, namely *integer aperture estimation*, is developed in chapter 5. It is shown that integer estimation in combination with the well-known ambiguity validation tests result in estimators, which belong to the class of integer aperture estimators. Furthermore, some alternative integer aperture estimators are presented, among them the optimal integer aperture estimator. It is defined such that the maximum success rate is obtained given a fixed fail rate.

Finally, the conclusions and recommendations are given in chapter 6.

2 GNSS observation model and quality control

In this chapter the GNSS observation model is presented. An observation model consists of the functional model, which describes the relation between the observations and the unknown parameters of interest, and of the stochastic model, which describes the stochastic properties of the observations. A great variety of GNSS models exists, which are used in applications like surveying, navigation and geophysics. An overview of these models can be found in textbooks like (Hofmann-Wellenhof et al. 2001; Leick 2003; Parkinson and Spilker 1996; Strang and Borre 1997; Teunissen and Kleusberg 1998).

Here, first a brief overview of the current and future Global Navigation Satellite Systems is given. In section 2.2 the GNSS observation equations are presented. Also, a description of the different error sources is given. With this information it is then possible to set up the functional model. Different types of models can be distinguished. Firstly, positioning and non-positioning models. Secondly, depending on the differencing technique that is applied.

The stochastic models corresponding to the different functional models, are presented in section 2.4. Finally, section 2.5 presents the general quality control theory applicable to GNSS models.

2.1 Global Navigation Satellite Systems

In the last decade the Global Positioning System (GPS) has found widespread use in all kind of applications. It is the first Global Navigation Satellite System (GNSS) offering the accuracy nowadays needed for e.g. surveying, navigation, and geophysics. More or less synchronously to GPS, the former Soviet Union has been developing a similar system under the name GLONASS. A third Global Navigation Satellite System is still under development and should be operational in 2010. It is the European Galileo system.

2.1.1 Global Positioning System

Currently, the nominal GPS constellation consists of 24 satellites, divided over six orbital planes, but more satellites are available at this moment (on average 28). The orbital inclination is 55^o and the orbital radius is 26,500 km.

Table 2.1: Signal and frequency plan for modernized GPS.

frequency band	frequency [MHz]	civil	precise	military
L1	1575.42	C/A	P	M
L2	1227.60	C/A	P	M
L5	1176.45	I + Q		

The principle of GPS is measuring distances between the GPS satellites and a receiver antenna. The ranging signals are transmitted on two carriers centered at the L1 and L2 frequency bands. Two types of Pseudo Random Noise (PRN) codes are modulated on these carriers. Only the C/A code on the L1 band is freely available for civil users. The P-codes are much more precise but are only available for military users; civil users can just access the encrypted P-codes, referred to as P(Y)-codes. The encryption of the P-codes is called Anti-Spoofing.

In the future a third frequency will be added, namely the L5 frequency. The codes modulated on the L5 carrier will be freely available. Furthermore, a civil code on L2 will become available, and military M-codes will be transmitted on L1 and L2. Table 2.1 gives an overview of the available signals.

2.1.2 GLONASS

The Russian equivalent to GPS is the GLObal NAvigation Satellite System (GLONASS). Like GPS, the nominal constellation consists of 24 satellites but with an orbital inclination of 64.8^o and an orbital radius of 25,500 km. GLONASS also transmits on the L1 and L2 frequency bands, but each satellite transmits its signals with a different frequency offset.

The complete GLONASS constellation of 24 operational satellites has only been available for a short time in 1996. Since then the number of satellites decreased and not enough replacements were launched. It is planned that GLONASS should become completely operational with a complete constellation in the period 2007-2022, and it is planned to transmit on a third frequency as well in the band 1190-1212 MHz, including integrity information.

2.1.3 Galileo

The Galileo satellite constellation will consist of 30 satellites, including three active spares. The satellites are divided over three orbital planes using the Walker 27/3/1 constellation plus the spares. The inclination will be 56^o and the orbital radius 29,600 km (Salgado et al. 2001; Zandbergen et al. 2004).

Galileo will transmit on four frequency bands. Table 2.2 gives an overview of the signals and the services for which they are available (Erhard 2002).

Table 2.2: Signal and frequency plan for Galileo. +: with integrity message; ++: with commercial data.

frequency band	frequency [MHz]	OS	SOL	CS	PRS
E2-L1-E1	1575.42	I/NAV	I/NAV +	I/NAV ++	G/NAV
E6	1278.75			C/NAV ++	G/NAV
E5b	1207.14	I/NAV	I/NAV +	I/NAV	
E5 (E5a+E5b)	1191.795				
E5a	1176.45	F/NAV	F/NAV	F/NAV	

The open service (OS) provides basic navigation and timing functions, using freely accessible signals and data, which are the Open Access Navigation Signals (F/NAV) on E5a, and the Integrity Navigation Signals (I/NAV) on E5b and E2-L1-E1 (also simply referred to as L1), but without the commercial data or integrity messages.

The Safety-of-Life Service (SOL) provides access to the same signals as OS users, but additionally SOL users have access to the I/NAV including the integrity messages transmitted on the E5b and E2-L1-E1 bands.

The Commercial Service (CS) is also available at the same signals as the OS, but users will have additional access to commercial data (for which they have to pay) transmitted on the Integrity Navigation Signals on E5b and E2-L1-E1, and to the Controlled Access Navigation Signals (C/NAV) on E6.

Finally, the Public Regulated Service (PRS) will include navigation and timing functions through Restricted Access Navigation Signals (G/NAV) on E6 and E2-L1-E1.

2.2 GNSS observation equations

Processing of the GNSS signals results in two basic observations: code and phase. The first is also referred to as pseudorange. The geometric distances between satellites and receivers can, however, not be derived from these observations directly. First, several bias terms have to be taken into account, as will be outlined in this section. Observation equations, which are parameterized in terms of geometric ranges, will be referred to as geometry-free observations equations. However, most users will not be interested in these distances, but in the (relative) receiver positions. Therefore, it will also be shown how the geometry-free observation equations must be linearized in order to arrive at geometry-based observation equations.

2.2.1 Code observations

The code or pseudorange observation is a coarse measure of the receiver-satellite distance. It is derived from the time difference between signal reception at receiver r and signal transmission at satellite s. The time of signal transmission is equal to the time

of reception diminished with the signal travel time. This gives the following basic code observation equation:

$$p^s_{r,j}(t) = c\,[t_r(t) - t^s(t - \tau^s_r)] + e^s_{r,j}(t) \tag{2.1}$$

with:

$p^s_{r,j}$	: code observation at receiver r from satellite s on frequency j [m]
t	: time of observation in GPS time [s]
c	: velocity of light in vacuum [m/s]
t_r	: reception time at receiver r [s]
t^s	: transmission time from satellite s [s]
τ	: signal travel time [s]
e	: code measurement error

Both the receiver clock time and satellite clock time are not exactly equal to GPS time. Therefore, the respective clock errors dt_r and dt^s need to be taken into account:

$$t_r(t) = t + dt_r(t) \tag{2.2}$$
$$t^s(t - \tau^s_r) = t - \tau^s_{r,j} + dt^s(t - \tau^s_r) \tag{2.3}$$

Inserting equations (2.2) and (2.3) in (2.1) yields

$$p^s_{r,j}(t) = c\tau^s_{r,j} + c\,[dt_r(t) - dt^s(t - \tau^s_r)] + e^s_{r,j}(t) \tag{2.4}$$

In order to determine the geometric distance between the satellite and the receiver, the signal travel time $\tau^s_{r,j}$ needs to be corrected for instrumental delays at the satellite and the receiver, and also for atmospheric effects and multipath:

$$\tau^s_{r,j} = \delta\tau^s_{r,j} + d_{r,j} + d^s_{,j} \tag{2.5}$$
$$\delta\tau^s_{r,j} = \frac{1}{c}\left[\rho^s_r + da^s_{r,j} + dm^s_{r,j}\right] \tag{2.6}$$

with:

$\delta\tau$	: signal travel time from satellite antenna to receiver antenna [s]
d_r	: instrumental code delay in receiver [s]
d^s	: instrumental code delay in satellite [s]
ρ	: geometric distance between satellite and receiver [m]
da	: atmospheric code error [m]
dm	: code multipath error [m]

Insertion of equations (2.5) and (2.6) into (2.4) gives the following code observation equation:

$$\begin{aligned} p^s_{r,j}(t) =& \rho^s_r(t, t - \tau^s_r) + da^s_{r,j}(t) + dm^s_{r,j}(t) \\ &+ c\left[dt_r(t) - dt^s(t - \tau^s_r) + d_{r,j}(t) + d^s_{,j}(t - \tau^s_r)\right] + e^s_{r,j}(t) \end{aligned} \tag{2.7}$$

2.2.2 Phase observations

The phase observation is a very precise but ambiguous measure of the geometric distance between a satellite and the receiver. The phase measurement equals the difference between the phase of the receiver-generated carrier signal at reception time, and the phase of the carrier signal generated in the satellite at transmission time. An integer number of full cycles is unknown since only the fractional phase is measured. This integer number is the so-called carrier phase ambiguity. The basic carrier phase observation equation is given by:

$$\varphi^s_{r,j}(t) = \varphi_{r,j}(t) - \varphi^s_{,j}(t - \tau^s_r) + N^s_{r,j} + \varepsilon^s_{r,j}(t) \tag{2.8}$$

with:

φ	: carrier phase observation [cycles]
N	: integer carrier phase ambiguity
ε	: phase measurement error

The phases on the right hand side are equal to:

$$\varphi_{r,j}(t) = f_j t_r(t) + \varphi_{r,j}(t_0) = f_j(t + dt_r(t)) + \varphi_{r,j}(t_0) \tag{2.9}$$

$$\varphi^s_{,j}(t) = f_j t^s(t - \tau^s_r) + \varphi^s_{,j}(t_0) = f_j(t - \tau^s_{r,j} + dt^s(t - \tau^s_r)) + \varphi^s_{,j}(t_0) \tag{2.10}$$

with:

f	: nominal carrier frequency [s^{-1}]
$\varphi_r(t_0)$	: initial phase in receiver at zero time [cycles]
$\varphi^s(t_0)$	: initial phase in satellite at zero time [cycles]

The carrier phase observation equation becomes:

$$\varphi^s_{r,j}(t) = f_j \left[\tau^s_{r,j} + dt_r(t) - dt^s(t - \tau^s_r)\right] + \left[\varphi_{r,j}(t_0) - \varphi^s_{,j}(t_0)\right] + N^s_{r,j} + \varepsilon^s_{r,j}(t) \tag{2.11}$$

This equation must be transformed to obtain units of meters and is therefore multiplied with the nominal wavelength of the carrier signal:

$$\phi_j = \lambda_j \varphi_j, \quad \text{with} \quad \lambda_j = \frac{c}{f_j}$$

The carrier signal travel time is expanded similarly as in equations (2.5) and (2.6). This results in the following observation equation:

$$\begin{aligned} \phi^s_{r,j}(t) =& \rho^s_r(t, t - \tau^s_r) + \delta a^s_{r,j}(t) + \delta m^s_{r,j}(t) \\ &+ c\left[dt_r(t) - dt^s(t - \tau^s_r) + \delta_{r,j}(t) + \delta^s_{,j}(t - \tau^s_r)\right] \\ &+ \left[\phi_{r,j}(t_0) + \phi^s_{,j}(t_0)\right] + \lambda_j N^s_{r,j} + \varepsilon^s_{r,j}(t) \end{aligned} \tag{2.12}$$

Note that the atmospheric delays, the multipath error, and the instrumental delays are different for code and phase measurements. Hence the different notation used in (2.12) (δ instead of d). Also note that the phase measurement error is multiplied with the wavelength, but the same notation ε is still used.

2.2.3 Signal travel time

In the preceding two sections the signal travel time τ_r^s in equations (2.7) and (2.12) has been assumed equal for code and phase measurements, as well as for measurements on different frequencies. In reality, this is not true since the signal travel time of the code actually is the group travel time. Moreover, the instrumental delays and atmospheric effects are different for the code and phase observations and for different frequencies. The difference between the travel times of different observation types and frequencies is, however, less than 10^{-7} s corresponding to sub-millimeter satellite position differences (Teunissen and Kleusberg 1998). Therefore, the travel times of all observation types are considered equal here.

2.2.4 Instrumental delays and clock errors

The receiver and satellite clock errors are independent of the observation type. On the other hand, the instrumental delays are different for code and phase measurements and for different frequencies. Still, the instrumental delays and clock errors will be lumped together because it is not possible to separate the effects in a least-squares adjustment. Hence:

$$\begin{array}{lcl} dt_{r,j}(t) & = & dt_r(t) + d_{r,j}(t) \\ dt^s_{,j}(t) & = & dt^s(t-\tau_r^s) - d^s_{,j}(t-\tau_r^s) \\ \delta t_{r,j}(t) & = & dt_r(t) + \delta_{r,j}(t) \\ \delta t^s_{,j}(t) & = & dt^s(t-\tau_r^s) - \delta^s_{,j}(t-\tau_r^s) \end{array} \tag{2.13}$$

2.2.5 Atmospheric delays

GNSS signals have to travel through the earth's atmosphere on their way from satellite to receiver. In the atmosphere neutral atoms and molecules, and charged particles are present that interact with the signals. This causes a propagation delay and signal absorption. The latter effect is not considered here. The propagation delay is caused by a deviation of the propagation speed from the velocity of light in vacuum, and an indirect delay due to signal bending. The delay is indicated with da and δa in equations (2.7) and (2.12) respectively.

The atmosphere can be divided in different "layers" with different characteristics. For GNSS especially the propagation characteristics are important, which means that the atmosphere needs to be divided into two "layers": the troposphere and the ionosphere.

Tropospheric delays

The troposphere is the lower part of the earth's atmosphere and ranges from the surface to a height of approximately 9 km at the poles and to 16 km at the equator. In the troposphere neutral atoms and molecules are present, which cause a signal delay. The troposphere is a non-dispersive medium at the GNSS frequencies, which means that

the propagation delay does not depend on the frequency of the signal. Moreover, the tropospheric delays are equal for carrier phase and code observations.

Usually a distinction is made between a wet delay caused by water vapor, and a hydrostatic delay caused by a mixture of dry air and water vapor that is considered to be in a hydrostatic equilibrium. The wet delay can be 0-40 cm in the zenith direction, and can only be determined with an accuracy of 2-5 cm based on semi-empirical models. The hydrostatic delay is much larger, 2.2-2.4 m in the zenith direction. However, it can be predicted with high accuracy based on surface pressure observations.

The tropospheric delay parameter will be denoted as T_r^s. A priori models can be used to correct for the tropospheric delays, see (Hopfield 1969; Saastamoinen 1973). It is necessary to estimate the tropospheric delay parameters if it is expected that one cannot fully rely on these models, see (Kleijer 2004).

Ionospheric delays

The ionized part of the atmosphere ranges approximately from 80 to 1000 km altitude. This atmospheric layer contains both uncharged and charged particles. The charged particles are created when neutral gas molecules are heated and electrons are liberated from them. This effect is called ionization. The rate of ionization depends on the density of the gas molecules and the intensity of the radiation. Because of the varying intensity of solar radiation, the free electron density is very variable in space and time.

The free electrons in the ionosphere affect the code and phase measurements differently. That is because the phase velocity is advanced, whereas the group velocity is delayed. So, the ionospheric effects on code and phase observations have opposite signs. Furthermore, the ionosphere is a dispersive medium, which means that the ionospheric effect is dependent on the frequency of the signal. The effects are inversely proportional to the square of the frequency of the signal. This allows for estimating the first order ionospheric effects by collecting measurements on different frequencies. Under worst-case conditions the first order term may be tens of meters. According to Odijk (2002) the higher order effects may be neglected in the case of relative positioning (see section 2.3.2) with inter-receiver distances up to 400 km.

The ionospheric effects on code and phase observations, respectively, are given by:

$$\begin{aligned} dI_{r,j}^s &= \mu_j I_r^s, \\ \delta I_{r,j}^s &= -\mu_j I_r^s, \end{aligned} \tag{2.14}$$

with

$$\mu_j = \frac{f_1^2}{f_j^2} = \frac{\lambda_j^2}{\lambda_1^2}. \tag{2.15}$$

The atmospheric effects in equations (2.7) and (2.12) can now be written as:

$$\begin{aligned} da_{r,j}^s &= T_r^s + \mu_j I_r^s \\ \delta a_{r,j}^s &= T_r^s - \mu_j I_r^s \end{aligned} \tag{2.16}$$

2.2.6 Multipath

Ideally, a GNSS signal should arrive at the receiver antenna only via the direct path from satellite to receiver. In practice, however, the signal often arrives at the antenna via two or more paths because the signal is reflected on nearby constructions or the ground. This effect is called multipath. It affects the phase and code measurements differently. The resulting measurement error depends on the strength of the direct and the reflected signals, and on the delay of the reflected signal.

Typically, the error induced by multipath on the code measurements varies between 1 and 5 meters. The effect on the phase measurements is 1-5 cm, and it will never be more than a quarter of a cycle provided that the amplitude of the direct signal is larger than the amplitude of the reflected signal.

Multipath is a systematic error and is very difficult to model. Although it is acknowledged that it is one of the most important issues in the field of high precision GNSS positioning, the multipath effect will be ignored here, i.e.

$$dm^s_{r,j} = \delta m^s_{r,j} = 0 \tag{2.17}$$

2.2.7 Random errors or noise

The GNSS code and phase observations are of a stochastic nature, so that random errors or observation noise must be taken into account in the observation equations. In equations (2.7) and (2.12) the noise was denoted as e and ε for code and phase respectively. The random errors are assumed to be zero-mean:

$$\begin{aligned} E\{e^s_{r,j}(t)\} &= 0 \\ E\{\varepsilon^s_{r,j}(t)\} &= 0 \end{aligned} \tag{2.18}$$

2.2.8 The geometry-free observation equations

The observation equations (2.7) and (2.12) are parameterized in terms of the satellite-receiver ranges ρ^s_r and are therefore referred to as non-positioning or geometry-free observation equations. The final non-positioning observation equations are obtained by inserting equations (2.13), (2.16), (2.17) into equations (2.7) and (2.12):

$$\begin{aligned} p^s_{r,j}(t) &= \rho^s_r(t) + T^s_r(t) + \mu_j I^s_r(t) + c dt_{r,j}(t) - c dt^s_{,j}(t) + e^s_{r,j}(t) \\ \phi^s_{r,j}(t) &= \rho^s_r(t) + T^s_r(t) - \mu_j I^s_r(t) + c\delta t_{r,j}(t) - c\delta t^s_{,j}(t) + \lambda_j M^s_{r,j} + \varepsilon^s_{r,j}(t) \end{aligned} \tag{2.19}$$

For notational convenience the receiver-satellite range is denoted as $\rho^s_r(t)$ instead of $\rho^s_r(t, t - \tau^s_r)$. Furthermore, the initial phases of the signal are lumped with the integer phase ambiguity, since these parameters are not separable. The resulting parameters in M^s_r are real-valued, i.e. non-integer:

$$M^s_{r,j} = \phi_{r,j}(t_0) + \phi^s_{,j}(t_0) + N^s_{r,j} \tag{2.20}$$

Here, they will still be referred to as ambiguities.

2.2.9 The geometry-based observation equations

The geometry-based observation equations are obtained after linearization of equations (2.19) with respect to the receiver position. The satellite-receiver range can be expressed as function of the satellite and receiver positions as follows:

$$\rho_r^s = \|\mathbf{r}^s - \mathbf{r}_r\| \tag{2.21}$$

where $\|\cdot\|$ indicates the length of a vector, $\mathbf{r}^s = \begin{bmatrix} x^s & y^s & z^s \end{bmatrix}^T$ the satellite position vector, and $\mathbf{r}_r = \begin{bmatrix} x_r & y_r & z_r \end{bmatrix}^T$ the receiver position vector.

In order to linearize the observation equations, approximate values of all parameters are needed. These approximate parameters will be denoted with a superscript ^{0}. It will be assumed that the approximate values of all parameters are zero, except for the satellite-receiver range, so that the observed-minus-computed code and phase observations are given by:

$$\begin{aligned} \Delta p_{r,j}^s(t) &= p_{r,j}^s(t) - \rho_r^{s,0}(t) \\ \Delta \phi_{r,j}^s(t) &= \phi_{r,j}^s(t) - \rho_r^{s,0}(t) \end{aligned} \tag{2.22}$$

The linearization of the increment $\Delta\rho_r^s$ is possible with the approximate satellite and receiver positions $\mathbf{r}^{s,0}$ and $\mathbf{r}_r^0$:

$$\Delta\rho_r^s = -(\mathbf{u}_r^s)^T \Delta\mathbf{r}_r + (\mathbf{u}_r^s)^T \Delta\mathbf{r}^s \tag{2.23}$$

with

$$\mathbf{u}_r^s = \frac{\mathbf{r}^{s,0} - \mathbf{r}_r^0}{\|\mathbf{r}^{s,0} - \mathbf{r}_r^0\|} \tag{2.24}$$

the unit line-of-sight (LOS) vector from receiver to satellite.

This gives the following linearized observation equations:

$$\begin{aligned} \Delta p_{r,j}^s(t) = &-(\mathbf{u}_r^s)^T \Delta\mathbf{r}_r + (\mathbf{u}_r^s)^T \Delta\mathbf{r}^s + T_r^s(t) + \mu_j I_r^s(t) \\ &+ cdt_{r,j}(t) - cdt_{,j}^s(t) + e_{r,j}^s(t) \\ \Delta \phi_{r,j}^s(t) = &-(\mathbf{u}_r^s)^T \Delta\mathbf{r}_r + (\mathbf{u}_r^s)^T \Delta\mathbf{r}^s + T_r^s(t) - \mu_j I_r^s(t) \\ &+ c\delta t_{r,j}(t) - c\delta t_{,j}^s(t) + \lambda_j M_{r,j}^s + \varepsilon_{r,j}^s(t) \end{aligned} \tag{2.25}$$

2.2.10 Satellite orbits

In order to compute approximate satellite-receiver ranges, approximate values of the satellite positions are required. Several sources are available in order to obtain these approximate values. Firstly, the information from the broadcast ephemerides can be used, which is available in real-time. Secondly, more precise information is made available by the International GPS Service (IGS). Finally, Yuma almanacs are freely available, e.g. from the website of the United States Coast Guard. These almanacs are especially interesting for design computations, but are not used for positioning. Table 2.3 gives an overview of the availability and accuracy of the different orbit products.

Table 2.3: Availability and accuracy of GPS satellite positions, after (Neilan et al. 2000).

Orbits	Availability	Accuracy
Yuma almanacs	Real-time or earlier	??
Broadcast	Real-time	5 m
IGS Ultra-rapid	Near real-time	20 cm
IGS Rapid	17 hours	10 cm
IGS Final	10 days	5 cm

2.3 GNSS functional model

In this section different GNSS functional models will be presented. Both the non-positioning or geometry-free models and the positioning or geometry-based models are presented.

Different GNSS models can also be distinguished based on the differencing that is applied. Differencing means taking the differences between observations from e.g. different receivers and/or different satellites. It is often applied in order to eliminate some of the parameters from the observation equations.

2.3.1 General mathematical model

The functional model describes the relationship between the observations and the unknown parameters. The m observation equations can be collected in the system $\underline{y} = Ax + \underline{e}$, where $\underline{y}$ and $\underline{e}$ are random, $\underline{e}$ is the discrepancy between $\underline{y}$ and Ax. It is assumed that the mean $E\{\underline{e}\}$ is zero, since $\underline{e}$ models the random nature of the variability in the measurements, and this variability will be zero 'on the average'. This gives for the expectation of $\underline{y}$:

$$E\{\underline{y}\} = Ax \tag{2.26}$$

The probability density function of $\underline{y}$ describes the variability in the outcomes of the measurements. For normally distributed data, it is completely captured by the dispersion:

$$D\{\underline{y}\} = E\{\underline{e}\underline{e}^T\} = Q_y \tag{2.27}$$

This is referred to as the stochastic model, with Q_y the variance-covariance (vc-) matrix of the observations. This matrix describes the precision of the observations and it is needed in order to properly weigh the observations in the adjustment process, see section 2.5. In section 2.4, the stochastic models corresponding to the functional models described in this section will be given.

The functional and stochastic model of (2.26) and (2.27) together are referred to as Gauss-Markov model.

2.3.2 Single difference models

If observations from m satellites are observed simultaneously at two receivers it is possible to take the differences of these observations. In this way, some parameters will be eliminated from the model. Here, only the model for a single epoch will be derived and the index t will be left out for notational convenience. The set of single difference (SD) observation equations for a single satellite, a single frequency, and two receivers q and r is then given by:

$$\begin{aligned} p^s_{qr,j} = p^s_{r,j} - p^s_{q,j} &= \rho^s_{qr} + T^s_{qr} + \mu_j I^s_{qr} + c dt_{qr,j} + e^s_{qr,j} \\ \phi^s_{qr,j} = \phi^s_{r,j} - \phi^s_{q,j} &= \rho^s_{qr} + T^s_{qr} - \mu_j I^s_{qr} + c\delta t_{qr,j} + \lambda_j M^s_{qr,j} + \varepsilon^s_{qr,j} \end{aligned} \tag{2.28}$$

with $(\cdot)_{qr} = (\cdot)_r - (\cdot)_q$.

Note that the lumped instrumental delays and clock errors of the satellite are eliminated. Furthermore, the initial carrier phase parameters of the satellite, which were contained in M^s_r and M^s_q, are cancelled out, so that M^s_{qr} only contains the differenced carrier phase ambiguity and the difference between the initial phases in the two receivers.

The observations of m satellites and two receivers observing on f frequencies will now be collected in the geometry-free functional model. The vector of observations becomes:

$$y = \begin{pmatrix} P \\ \Phi \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} \left(p^1_{qr,1} & \cdots & p^m_{qr,1}\right) & \cdots & \left(p^1_{qr,f} & \cdots & p^m_{qr,f}\right) \end{bmatrix}^T \\ \begin{bmatrix} \left(\phi^1_{qr,1} & \cdots & \phi^m_{qr,1}\right) & \cdots & \left(\phi^1_{qr,f} & \cdots & \phi^m_{qr,f}\right) \end{bmatrix}^T \end{pmatrix} \tag{2.29}$$

The parameter vectors are defined as:

$$\rho = \begin{pmatrix} \rho^1_{qr} & \cdots & \rho^m_{qr} \end{pmatrix}^T \tag{2.30}$$

$$T = \begin{pmatrix} T^1_{qr} & \cdots & T^m_{qr} \end{pmatrix}^T \tag{2.31}$$

$$\mathcal{I} = \begin{pmatrix} I^1_{qr} & \cdots & I^m_{qr} \end{pmatrix}^T \tag{2.32}$$

$$\mathbf{dt}_{qr} = c \cdot \begin{pmatrix} \begin{bmatrix} dt_{qr,1} & \cdots & dt_{qr,f} \end{bmatrix}^T \\ \begin{bmatrix} \delta t_{qr,1} & \cdots & \delta t_{qr,f} \end{bmatrix}^T \end{pmatrix} \tag{2.33}$$

$$a = \begin{pmatrix} \begin{bmatrix} M^1_{qr,1} & \cdots & M^m_{qr,1} \end{bmatrix} & \cdots & \begin{bmatrix} M^1_{qr,f} & \cdots & M^m_{qr,f} \end{bmatrix} \end{pmatrix}^T \tag{2.34}$$

For notational convenience the Kronecker product is used, see appendix A.1, to arrive at the geometry-free model:

$$\begin{aligned} E\{\underline{y}\} = &\left(e_{2f} \otimes I_m\right) \rho + \left(e_{2f} \otimes I_m\right) T + \left(\begin{bmatrix} \mu \\ -\mu \end{bmatrix} \otimes I_m \right) \mathcal{I} \\ &+ \left(I_{2f} \otimes e_m\right) \mathbf{dt}_{qr} + \left(\begin{bmatrix} 0 \\ \Lambda \end{bmatrix} \otimes I_m \right) a \end{aligned} \tag{2.35}$$

with e_p a p-vector with ones, I_p an identity matrix of order p, and:

$$\mu = \begin{pmatrix} \mu_1 & \cdots & \mu_f \end{pmatrix}^T, \qquad \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_f \end{pmatrix} \tag{2.36}$$

In the same way the geometry-based model is obtained, but the observations collected in the vector y are now the observed-minus-computed observations from equation (2.22). The model is given by:

$$\begin{aligned} E\{\underline{y}\} = &\left(e_{2f} \otimes G\right) \Delta\mathbf{r}_{qr} + \left(e_{2f} \otimes I_m\right) T + \left(\begin{bmatrix} \mu \\ -\mu \end{bmatrix} \otimes I_m\right) \mathcal{I} \\ &+ \left(I_{2f} \otimes e_m\right) \mathbf{dt}_{qr} + \left(\begin{bmatrix} 0 \\ \Lambda \end{bmatrix} \otimes I_m\right) a \end{aligned} \tag{2.37}$$

with

$$G = \begin{bmatrix} -\mathbf{u}_{qr}^1 & \cdots & -\mathbf{u}_{qr}^m \end{bmatrix}^T \tag{2.38}$$

and

$$\begin{aligned} \left(-(\mathbf{u}_r^s)^T \Delta\mathbf{r}_r + (\mathbf{u}_r^s)^T \Delta\mathbf{r}^s\right) - &\left(-(\mathbf{u}_q^s)^T \Delta\mathbf{r}_q + (\mathbf{u}_q^s)^T \Delta\mathbf{r}^s\right) \\ &= -(\mathbf{u}_{qr}^s)^T \Delta\mathbf{r}_{qr} - (\mathbf{u}_{qr}^s)^T \Delta\mathbf{r}_q + (\mathbf{u}_{qr}^s)^T \Delta\mathbf{r}^s \\ &\approx -(\mathbf{u}_{qr}^s)^T \Delta\mathbf{r}_{qr}, \end{aligned} \tag{2.39}$$

since the orbital uncertainty $\Delta\mathbf{r}^s$ can be ignored if the baseline length $\|\mathbf{r}_{qr}\|$ is small compared to the satellite altitude (Teunissen and Kleusberg 1998). The term with the receiver position error $\Delta\mathbf{r}_q$ is zero when the position of receiver q is assumed known. The parameters in $\Delta\mathbf{r}_{qr}$ are referred to as the baseline increments; if the receiver position $\mathbf{r}_q$ is known, the unknown receiver position $\mathbf{r}_q$ can be determined from the estimated $\Delta\mathbf{r}_{qr}$.

Both single difference models (2.35) and (2.37) are rank deficient. The rank deficiencies can be resolved by lumping some of the parameters and then estimate these lumped parameters instead of the original ones. The parameters are lumped by performing a parameter transformation, which corresponds to the so-called *S-basis* technique as developed by Baarda (1973) and Teunissen (1984).

A first rank deficiency is caused by the inclusion of the tropospheric parameters. This can be solved for the geometry-free model by lumping the tropospheric parameters with the ranges:

$$\rho = \begin{pmatrix} \rho_{qr}^1 + T_{qr}^1 & \cdots & \rho_{qr}^m + T_{qr}^m \end{pmatrix}^T \tag{2.40}$$

In the geometry-based case, the tropospheric delay parameter is first reparameterized in a dry and wet component. The dry component is responsible for about 90% of the tropospheric delay and can be reasonably well corrected for by using a priori tropospheric models, e.g. (Saastamoinen 1973). The much smaller wet component is quite variable

and difficult to predict. Therefore, the wet delays are commonly mapped to zenith tropospheric delay (ZTD) parameters and these are then estimated. An overview of available mapping functions can be found in Kleijer (2004).

For short time spans this ZTD parameter can be considered constant. The parameter vector T consists then of the ZTD parameter of one receiver. The mapping function is denoted as ψ_r^s. This gives for the partial design matrix related to the ZTD parameter:

$$e_{2f} \otimes \begin{bmatrix} \psi_r^1 & \cdots & \psi_r^m \end{bmatrix}^T = e_{2f} \otimes \Psi \tag{2.41}$$

Note that the a priori tropospheric corrections and the ZTD of the reference station q are added to the approximate observation equations in (2.22).

the ZTD is not estimated if it is assumed that the a priori model can be fully relied on. This is referred to as the troposphere-fixed approach. If, on the other hand, the ZTD is estimated, this is referred to as the troposphere-float approach.

The rank deficiency caused by the phase receiver clocks δt_r^s and the ambiguities is solved by the following transformations:

$$\mathbf{dt}_{qr} = \begin{pmatrix} \begin{bmatrix} cdt_{qr,1} + \lambda_1 M_{qr,1}^1 & \cdots & cdt_{qr,f} + \lambda_f M_{qr,f}^1 \end{bmatrix}^T \\ \begin{bmatrix} c\delta t_{qr,1} + \lambda_1 M_{qr,1}^1 & \cdots & c\delta t_{qr,f} + \lambda_f M_{qr,f}^1 \end{bmatrix}^T \end{pmatrix} \tag{2.42}$$

$$a = \begin{pmatrix} \begin{bmatrix} N_{qr,1}^{12} & \cdots & N_{qr,1}^{1m} \end{bmatrix} & \cdots & \begin{bmatrix} N_{qr,f}^{12} & \cdots & N_{qr,f}^{1m} \end{bmatrix} \end{pmatrix}^T \tag{2.43}$$

Due to this reparameterization, the single difference ambiguities are transformed to double difference ambiguities from which it is known that they are of integer nature since the initial phases in receiver and satellite are cancelled out: $M_{qr,j}^{ts} = N_{qr,j}^{ts}$. Note that the code receiver clock parameters are not transformed.

Finally, a rank deficiency is present due to the inclusion of the ionospheric parameters. In Odijk (2002) it is described how to set up the so-called ionosphere-weighted model. The approach is to include a vector of ionospheric pseudo-observations, consisting of a priori estimates, to the vector of observations:

$$y = \begin{pmatrix} P \\ \Phi \\ \mathcal{I} \end{pmatrix} \tag{2.44}$$

If the a priori information is considered exact, the a priori estimates can simply be subtracted from the observations and there are no ionospheric parameters to be estimated. This is referred to as the ionosphere-fixed model. If the baseline is shorter than 10 km, it can be assumed that the ionospheric delay on the signal of one satellite to the two receivers is identical, i.e. $I_r^s = I_q^s$. Also then the ionospheric parameters can be removed from the single difference model. If, on the other hand, the ionospheric behavior must be considered completely unknown, e.g. when the baseline is very long, the weight of the ionospheric pseudo-observations is set equal to zero, which is equivalent to setting the standard deviation of the pseudo-observations to infinity. This is referred to as the ionosphere-float model.

The single difference models can now be presented as follows. The full-rank geometry-free model is given by:

$$\begin{aligned} E\{\underline{y}\} = &\left(\begin{bmatrix} e_{2f} \\ 0 \end{bmatrix} \otimes I_m\right) \rho + \left(\begin{bmatrix} \mu \\ -\mu \\ 1 \end{bmatrix} \otimes I_m\right) \mathcal{I} \\ &+ \left(\begin{bmatrix} I_{2f} \\ 0 \end{bmatrix} \otimes e_m\right) \mathbf{dt}_{qr} + \left(\begin{bmatrix} 0 \\ \Lambda \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix}\right) a \end{aligned} \tag{2.45}$$

The observation vector is given by (2.44), the range vector ρ by (2.40), and the parameter vectors $\mathcal{I}$, $\mathbf{dt}_{qr}$, and a by equations (2.32), (2.42) and (2.43) respectively.

The ionosphere-weighted, troposphere-float, geometry-based model is obtained as:

$$\begin{aligned} E\{\underline{y}\} = &\left(\begin{bmatrix} e_{2f} \\ 0 \end{bmatrix} \otimes G\right) \Delta\mathbf{r}_{qr} + \left(\begin{bmatrix} e_{2f} \\ 0 \end{bmatrix} \otimes \Psi\right) T + \left(\begin{bmatrix} \mu \\ -\mu \\ 1 \end{bmatrix} \otimes I_m\right) \mathcal{I} \\ &+ \left(\begin{bmatrix} I_{2f} \\ 0 \end{bmatrix} \otimes e_m\right) \mathbf{dt}_{qr} + \left(\begin{bmatrix} 0 \\ \Lambda \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix}\right) a \end{aligned} \tag{2.46}$$

2.3.3 Double difference models

The number of parameters in the observation equations can be even further reduced by also taking differences between observations of different satellites. In this double difference (DD) approach, one reference satellite t is chosen and the single difference observations of this satellite are subtracted from the corresponding single difference observations of all other satellites. For one satellite-pair the geometry-free observation equations become then:

$$\begin{aligned} p^{ts}_{qr,j} = p^{s}_{qr,j} - p^{t}_{qr,j} &= \rho^{ts}_{qr} + T^{ts}_{qr} + \mu_j I^{ts}_{qr} + e^{s}_{qr,j} \\ \phi^{ts}_{qr,j} = \phi^{s}_{qr,j} - \phi^{t}_{qr,j} &= \rho^{ts}_{qr} + T^{ts}_{qr} - \mu_j I^{ts}_{qr} + \lambda_j N^{ts}_{qr,j} + \varepsilon^{ts}_{qr,j} \end{aligned} \tag{2.47}$$

The instrumental delays and clock errors of the receivers have now also cancelled from the observation equations, as well as the initial phases in the receivers. This leaves only the integer carrier phase ambiguity as extra parameter in the phase observations. Only if this ambiguity is resolved, the phase observations can be considered as very precise pseudorange measurements, so that high precision positioning solutions can be obtained. Therefore, the integer nature of the ambiguities must be exploited. Integer parameter resolution is, however, a non-trivial problem, which will be the topic of the remaining chapters.

If the first satellite is chosen as the reference satellite, the following transformation matrix can be used to arrive at the double difference geometry-free model:

$$D^T = \begin{pmatrix} -e_{m-1} & I_{m-1} \end{pmatrix} \tag{2.48}$$

The double difference observation vector and parameter vectors are now derived from the single difference vectors (denoted with superscript sd) of equations (2.29)-(2.34) as follows:

$$y = \left(I_{(2f+1)} \otimes D^T\right) y^{\text{sd}} \tag{2.49}$$
$$\rho = D^T \rho^{\text{sd}} \tag{2.50}$$
$$T = D^T T^{\text{sd}} \tag{2.51}$$
$$\mathcal{I} = D^T \mathcal{I}^{\text{sd}} \tag{2.52}$$
$$a = \left(I_f \otimes D^T\right) a^{\text{sd}} \tag{2.53}$$

This gives the ionosphere-weighted double difference geometry-free model:

$$E\{\underline{y}\} = \left(\begin{bmatrix} e_{2f} \\ 0 \end{bmatrix} \otimes I_{m-1}\right) \rho + \left(\begin{bmatrix} \mu \\ -\mu \\ 1 \end{bmatrix} \otimes I_{m-1}\right) \mathcal{I} + \left(\begin{bmatrix} 0 \\ \Lambda \\ 0 \end{bmatrix} \otimes I_{m-1}\right) a \tag{2.54}$$

with ρ the lumped tropospheric and range parameters.

Similarly, the geometry-based model is obtained as:

$$\begin{aligned} E\{\underline{y}\} = &\left(\begin{bmatrix} e_{2f} \\ 0 \end{bmatrix} \otimes G\right) \Delta\mathbf{r}_{qr} + \left(\begin{bmatrix} e_{2f} \\ 0 \end{bmatrix} \otimes \Psi\right) T + \left(\begin{bmatrix} \mu \\ -\mu \\ 1 \end{bmatrix} \otimes I_{m-1}\right) \mathcal{I} \\ &+ \left(\begin{bmatrix} 0 \\ \Lambda \\ 0 \end{bmatrix} \otimes I_{m-1}\right) a \end{aligned} \tag{2.55}$$

with $G = D^T G^{\text{sd}}$, and T contains the ZTD.

The double difference approach offers the advantage of less (or even no) rank deficiencies in the mathematical model. However, rank deficiencies can also be resolved in other ways as explained in the preceding section.

A disadvantage is that users will be interested in the undifferenced parameters, and therefore one has to be careful with the interpretation of the results.

As double differencing eliminates the satellite and receiver clock errors, it is not possible to estimate them explicitly. This implies that it is not possible to model their behavior in time. But these parameters are rarely needed, so in many practical situations this will not be considered as a problem.

An important disadvantage of the double difference approach is that all receivers must track the same satellites. In practice this is not always the case, especially with long baselines. But also when the receivers are located quite close to each other, one of the receivers may not track some of the satellites tracked by the other receiver(s), e.g. due to blocking of signals by local obstacles. In such cases using the double difference model results in loss of information, as observations from a satellite only tracked by one of the receivers cannot be used.

The single difference approach can be seen as a kind of compromise. The amount of observations is reduced with respect to the undifferenced approach, but not as much as with the double difference approach. The same holds for the unknown parameters. An advantage of the single difference approach is that in the case of a single baseline model (two receivers) no correlation between the observations is introduced, see section 2.4.

As explained before, for high-precision applications it is important that the carrier phase ambiguities are resolved as integers. It has been shown that the parameterization of the unknown ambiguities is such that these parameters are known to be integers in the single difference and double difference approach.

2.4 GNSS stochastic model

The stochastic model describes the precision of the observations. As a starting point the following assumptions are made:

$$\begin{aligned} E\{\underline{e}^s_{r,j}(t)\underline{e}^s_{r,j}(t)\} &= \sigma^2_{p,j} && \text{no satellite dependent weighting} \\ E\{\underline{e}^s_{r,i}(t)\underline{e}^s_{r,j}(t)\} &= 0 && \text{no frequency dependent correlation} \\ E\{\underline{e}^s_{r,j}(t_1)\underline{e}^s_{r,j}(t_2)\} &= 0 && \text{no time correlation} \end{aligned} \tag{2.56}$$

Often, the code standard deviation is chosen identical for all frequencies, $\sigma_{p,j} = \sigma_p \, \forall j$. The same assumptions are made for the undifferenced phase observations, with the phase standard deviation denoted as σ_ϕ.

The vc-matrix of the undifferenced observations of one satellite and one receiver becomes then:

$$C_{p\phi} = diag(\sigma^2_{p,1}, \ldots, \sigma^2_{p,f}, \sigma^2_{\phi,1}, \ldots, \sigma^2_{\phi,f}, \sigma^2_I) \tag{2.57}$$

where σ^2_I is the variance of the ionospheric pseudo-observation, which has to be included if the ionosphere-weighted model is considered.

For the single differenced observations, the vc-matrices of the observations of both receivers must be added according to the propagation law of variances, see appendix A.2.1:

$$C^{\mathsf{sd}}_{p\phi} = C_{p\phi,q} + C_{p\phi,r} \tag{2.58}$$

If the standard deviations are assumed equal for both receivers, all variances must simply be multiplied by 2.

The single difference vc-matrix corresponding to the functional models in equations (2.45) and (2.46) becomes:

$$Q_y = C^{\mathsf{sd}}_{p\phi} \otimes I_m \tag{2.59}$$

The vc-matrix of the double difference models in equations (2.54) and (2.55) is obtained again with the propagation law of variances and the properties of the Kronecker product

(appendix A.1):

$$Q_y = C_{p\phi}^{\mathsf{sd}} \otimes D^T D \tag{2.60}$$

with the transformation matrix D^T as defined in (2.48).

2.4.1 Variance component estimation

The stochastic model as specified by the vc-matrix of the observations can be given as:

$$Q_y = \sigma^2 G_y \tag{2.61}$$

with σ^2 the variance factor of unit weight, and G_y the cofactor matrix. If the variance factor is assumed to be unknown, it can be estimated a posteriori. For that purpose equation (2.61) can be generalized into:

$$Q_y = \sum_{\alpha}^{p} \sigma_\alpha^2 G_\alpha \tag{2.62}$$

Hence, the stochastic model is written as a linear combination of unknown factors and known cofactor matrices, (Tiberius and Kenselaar 2000). The factors σ_α^2 can represent variances as well as covariances, for example variance factors for the code and phase data, a variance for each satellite-receiver pair, and the covariance between code and phase data. The problem of estimating the unknown (co-)variance factors is referred to as variance component estimation.

2.4.2 Elevation dependency

Until now the standard deviations of a certain observation type were chosen equal for all satellites, although it might be more realistic to assign weights to the observations of different satellites depending on their elevation. That is because increased signal attenuation due to the atmosphere and multipath effects will result in less precise observations from satellites at low elevations. Therefore, Euler and Goad (1991) have suggested to use elevation-dependent weighting:

$$\sigma_{p_r^s}^2 = (q_r^s)^2 \sigma_p^2 \tag{2.63}$$

with σ_p the standard deviation of code observations to satellites in the zenith, and

$$q_r^s = 1 + a \cdot \exp\{\frac{-\varepsilon_r^s}{\varepsilon_0}\}, \tag{2.64}$$

where a is an amplification factor, ε_r^s the elevation angle, and ε_0 a reference elevation angle. The values of a and ε_0 depend on the receiver that is used, and on the observation type – the function q in equation (2.64) may be different for code and phase observations as well as for observations on different frequencies.

Experimental results have shown that an improved stochastic model can be obtained if elevation dependent weighting is applied, see e.g. (Liu 2002).

2.4.3 Cross-correlation and time correlation

In equation (2.56) the assumption was made that there is no frequency dependent correlation. Several studies have shown, however, that this assumption may not be realistic, see e.g. (Bona 2000; Bona and Tiberius 2001; Liu 2002). It was shown that the stochastic model can be refined by performing receiver tests or variance component estimation, in order to determine the cross-correlation.

It was also assumed that no time correlation is present, although the observations collected at different epochs may be correlated when the sampling frequency is high. In Borre and Tiberius (2000) significant time correlation was found for 5 Hz data. If, on the other hand, the sampling interval is larger than 30 s, according to Bona (2000) time correlation can be neglected. So, time correlation does not play a role for instantaneous applications.

2.5 Least-squares estimation and quality control

In sections 2.3 and s:stochastic, the GNSS functional models and stochastic models were presented. In order to solve the observation model, a least-squares adjustment is generally used, since it is simple to apply and, by definition, it has important optimality properties. For a model written as equation (2.26) the least-squares estimator is defined as:

$$\underline{\hat{x}} = \arg\min_{x} \|\underline{y} - Ax\|^2_{Q_y} \tag{2.65}$$

and its solution is given by:

$$\begin{aligned} \underline{\hat{x}} &= (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} \underline{y} \\ \underline{\hat{y}} &= A\underline{\hat{x}} \\ \underline{\hat{e}} &= \underline{\hat{y}} - \underline{y} \end{aligned} \tag{2.66}$$

The GNSS observations are generally considered to be normally distributed, and then least-squares estimation is equivalent to maximum-likelihood estimation, and to best linear unbiased estimation if the model is linear, (Teunissen 2000a).

The estimator $\underline{\hat{x}}$ is a random vector since it is functionally related to the random vector $\underline{y}$. Because of the linear relationship between the two vectors, $\underline{\hat{x}}$ will also have a normal distribution like $\underline{y}$. Therefore the distribution of the estimator can be characterized by its mean and dispersion, which are given as:

$$E\{\underline{\hat{x}}\} = x \tag{2.67}$$

$$D\{\underline{\hat{x}}\} = Q_{\hat{x}} = (A^T Q_y^{-1} A)^{-1} \tag{2.68}$$

This shows that the estimator is unbiased, i.e. that the expectation of the estimator equals the unknown and sought for parameter vector x. The vc-matrix describes the precision of the estimator, which is independent of the observations, so that the precision is known without the need for actual measurements.

The quality of the solution can then be measured by the precision of the estimators. However, the precision does not give information on the validity of the model, so that the unbiasedness of the estimators cannot be guaranteed. It is therefore important to use statistical tests in order to get information on the validity of the model.

2.5.1 Model testing

Statistical testing is possible if there is a null hypothesis, H_0, that can be tested against an alternative hypothesis, H_a. These hypotheses can be defined as:

$$\begin{aligned} H_0 &: E\{\underline{y}\} = Ax; \qquad D\{\underline{y}\} = Q_y \\ H_a &: E\{\underline{y}\} = Ax + C\nabla; \qquad D\{\underline{y}\} = Q_y \end{aligned} \tag{2.69}$$

where C is a known $m \times q$-matrix that specifies the type of model error, and ∇ an unknown q-vector. It is assumed that $\underline{y}$ is normally distributed, see appendix A.2.

H_0 can be tested against H_a using the following test statistic:

$$\underline{T}_q = \frac{1}{q}\underline{\hat{e}}^T Q_y^{-1} C \left(C^T Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C\right)^{-1} C^T Q_y^{-1} \underline{\hat{e}} \tag{2.70}$$

The test statistic $\underline{T}_q$ has a central F-distribution with q and ∞ degrees of freedom under H_0, and a non-central F-distribution under H_a:

$$H_0 : \underline{T}_q \sim F(q, \infty, 0); \quad H_a : \underline{T}_q \sim F(q, \infty, \lambda) \tag{2.71}$$

with non-centrality parameter λ:

$$\lambda = \nabla^T C^T Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C \nabla \tag{2.72}$$

The test is then given by:

$$\text{reject } H_0 \text{ if } \quad \underline{T}_q > F_\alpha(q, \infty, 0), \tag{2.73}$$

where α is a chosen value of the level of significance, also referred to as the false alarm rate since it equals the probability of rejecting H_0 when in fact it is true. $F_\alpha(q, \infty, 0)$ is the critical value such that:

$$\alpha = \int_{F_\alpha(q,\infty,0)}^{\infty} f_{\underline{F}}(F|q, \infty, 0) dF \tag{2.74}$$

where $f_{\underline{F}}(F|q, \infty, 0)$ is the probability density function of $F(q, \infty, 0)$, see appendix A.2.3.

The value of $\lambda = \lambda_0$ can be computed once reference values are known for the level of significance $\alpha = \alpha_q$, and the detection power $\gamma = \gamma_0$, which is the probability of rejecting H_0 when indeed H_a is true:

$$\gamma = \int_{F_\alpha(q,\infty,0)}^{\infty} f_{\underline{F}}(F|q, \infty, \lambda) dF \tag{2.75}$$

So, $\lambda_0 = \lambda(\alpha_q, \gamma_0, q)$. From this value one can then compute the corresponding model error $C\nabla$ from equation (2.72).

The vector $\nabla y = C\nabla$ describes the *internal reliability* of hypothesis H_0 with respect to H_a. The internal reliability is thus a measure of the model error that can be detected with a probability $\gamma = \gamma_0$ by test (2.73). This measure can be computed by considering certain types of model errors, as specified by the matrix C. In GNSS applications common error types are for example carrier cycle slips and code outliers. In those cases, $q = 1$, so that C reduces to a vector c, and ∇ to a scalar. From equation (2.72) follows then that:

$$|\nabla| = \sqrt{\frac{\lambda_0}{c^T Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c}} \tag{2.76}$$

$|\nabla|$ is called the *minimal detectable bias* (MDB). With $|\nabla|$ known, it is possible to determine the corresponding impact of the model error on the estimates of the unknown parameters, referred to as the *external reliability*. This impact can be computed as:

$$\nabla\hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} \nabla y \tag{2.77}$$

by using $\nabla y = c|\nabla|$. $\nabla\hat{x}$ is referred to as the *minimal detectable effect* (MDE).

Note that the measures of the internal reliability and the external reliability can both be computed without the need for actual observations, and hence can be used for planning purposes before the measurements are carried out.

2.5.2 Detection, Identification and Adaptation

Once the observations are collected there will be several alternative hypotheses that need to be tested against the null hypothesis. The DIA-procedure (detection, identification, adaptation) can be used in order to have a structured testing procedure (Teunissen 1990).

The first step in DIA is *detection*, for which an overall model test is performed to diagnose whether an unspecified model error has occurred. In this step the null hypothesis is tested against the most relaxed alternative hypothesis:

$$H_a : E\{\underline{y}\} \in \mathbb{R}^m \tag{2.78}$$

The appropriate test statistic for detection is then:

$$\underline{T}_{m-n} = \frac{\underline{\hat{e}}^T Q_y^{-1} \underline{\hat{e}}}{m-n} \tag{2.79}$$

The null hypothesis will be rejected if $\underline{T}_{m-n} > F_{\alpha_{m-n}}(m-n, \infty, 0)$. In that case, the next step will be the *identification* of the model error.

The most likely alternative hypothesis is the one for which

$$\frac{\underline{T}_{q_i}}{F_{\alpha_{q_i}}(q_i, \infty, 0)} \tag{2.80}$$

is maximum. The degrees of freedom q_i may be different, and the level of significance α_{q_i} depends on $\lambda(\alpha_{q_i}, \gamma, q_i)$, which is chosen as:

$$\lambda(\alpha_{q_i}, \gamma, q_i) = \lambda(\alpha_{m-n}, \gamma, m-n) \tag{2.81}$$

This means that the non-centrality parameters and detection powers are chosen equal for both the detection and the identification test, implying that a certain model error can be found with the same probability by both tests. This approach is the *B-method of testing*, (Baarda 1968).

In practice the identification step is started with the so-called *data snooping*, (Baarda 1968), which means that each individual observation is screened for the presence of an outlier. In that case $q = 1$ and the matrix C reduces to a vector, denoted as c. The test statistic of equation (2.70) becomes:

$$\begin{aligned} \underline{T}_1 &= (\underline{w})^2 \\ \underline{w} &= \frac{c^T Q_y^{-1} \underline{\hat{e}}}{\sqrt{c^T Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c}} \end{aligned} \tag{2.82}$$

The test statistic $\underline{w}$ has a standard normal distribution $N(0,1)$ under H_0. An observation j is suspected to be biased if:

$$|w_j| \geq |w_i| \, \forall i \quad \text{and} \quad |w_j| > N_{\frac{1}{2}\alpha}(0,1) \tag{2.83}$$

The final step in the DIA-procedure is *adaptation*, in which the model is corrected for one or more of the most likely model errors identified in the previous step. This may involve replacement of the erroneous data, e.g. by re-measurement, or a new null hypothesis is set up which takes into account the identified model errors. After this step, one has to make sure that the new situation will lead to acceptance of the null hypothesis, meaning that the DIA-procedure needs to be applied iteratively.

2.5.3 GNSS quality control

In this section a general testing procedure was described that can be applied in the case of linear estimation with normally distributed data, so that the parameter distribution is completely captured by the vc-matrix of the estimators. Unfortunately, this approach cannot be applied when integer parameters are involved in the estimation process, since these parameters do not have a Gaussian distribution, and therefore the parameter distribution itself is needed to obtain appropriate test statistics in order to validate the integer solution. This problem is the subject of section 3.5 and chapter 5.

3 Integer ambiguity resolution

The problem of integer ambiguity resolution has drawn a lot of attention in the past decades. Many ambiguity resolution algorithms have been published. Not all of these algorithms will be discussed here; only a brief overview will be given. The first step of ambiguity resolution is integer estimation. Therefore, section 3.1 starts with the definition of admissible integer estimators. For the quality description and validation of the estimators, their distribution functions are required. The distributional properties of the float and fixed ambiguity will be given in section 3.2, where it is also shown how the probability of correct integer estimation, the success rate, can be approximated. For the purpose of validation, the parameter distribution of the ambiguity residuals is required. This distribution function and its properties will be given in section 3.3. In section 3.4 it is shown how the quality of the fixed baseline estimator can be expressed. Section 3.5 gives an overview of currently available methods for the validation of the fixed ambiguity solution and their shortcomings. Finally, in section 3.6 a completely different approach of ambiguity resolution is described, namely the Bayesian approach.

3.1 Integer estimation

Any GNSS observation model can be parameterized in integers and non-integers. This gives the following system of linear(ized) observation equations:

$$\underset{m\times 1}{\underline{y}} = \underset{m\times nn\times 1}{A\ a} + \underset{m\times pp\times 1}{B\ b} + \underset{m\times 1}{\underline{e}} \tag{3.1}$$

where y is the GPS observation vector of order m, a and b are the unknown parameter vectors of dimension n and p respectively, and $\underline{e}$ is the noise vector. The data vector y usually consists of the observed-minus-computed DD phase and/or code observations on one, two or three frequencies and accumulated over all observation epochs. The entries of the parameter vector a will then consist of the unknown integer carrier phase ambiguities, which are expressed in units of cycles rather than in units of range. It is known that the entries are integers, so that $a \in \mathbb{Z}^n$. The remaining unknown parameters form the entries of the vector b. These parameters may be the unknown baseline increments and for instance atmospheric (ionospheric, tropospheric) delays, which are all real-valued, i.e. $b \in \mathbb{R}^p$. These real-valued parameters are referred to as the baseline parameters, although the vector b may thus contain other parameters than only the baseline components.

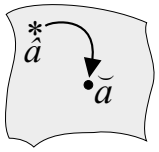

Figure 3.1: An ambiguity pull-in region of $z = \breve{a}$.

The classical linear estimation theory can be applied to models that contain real-valued parameters. However, if the integerness of the ambiguity parameters is taken into account, a different approach must be followed which includes a separate step for ambiguity resolution. The complete estimation process will then consist of three steps (Teunissen 1993). In the first step, the integerness of the vector a is discarded and the so-called float solution is computed with a standard least-squares adjustment. This results in real-valued estimates for a and b and their vc-matrix:

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}; \quad \begin{pmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{a}\hat{b}} & Q_{\hat{b}} \end{pmatrix} \tag{3.2}$$

In the second step the integer ambiguity estimate is computed from the float ambiguity estimate $\hat{a}$:

$$\breve{a} = S(\hat{a}) \tag{3.3}$$

where $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$ is the mapping from the n-dimensional space of real numbers to the n-dimensional space of integers. The final step is to use the integer ambiguity estimates to correct the float estimate of b with

$$\breve{b} = \hat{b}(\breve{a}) = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \breve{a}) \tag{3.4}$$

This solution is referred to as the fixed baseline solution. Equations (3.3) and (3.4) depend on the choice of the integer estimator. Different integer estimators are obtained for different choices of the map $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$. This implies that also the probability distribution of the estimators depends on the choice of the map.

In order to arrive at a class of integer estimators, first the map $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$ will be considered. The space of integers, $\mathbb{Z}^n$, is of a discrete nature, which implies that the map must be a many-to-one map, and not one-to-one. In other words, different real-valued ambiguity vectors a will be mapped to the same integer vector. Therefore, a subset $S_z \subset \mathbb{R}^n$ can be assigned to each integer vector $z \in \mathbb{Z}^n$:

$$S_z = \{x \in \mathbb{R}^n \mid \quad z = S(x)\}, \quad z \in \mathbb{Z}^n \tag{3.5}$$

This subset S_z contains all real-valued float ambiguity vectors that will be mapped to the same integer vector z, and it is called the pull-in region of z (Jonkman 1998; Teunissen 1998c), see figure 3.1. This implies that $\breve{a} = z \Leftrightarrow \hat{a} \in S_z$. The integer ambiguity estimator can be expressed as[1]:

$$\breve{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a}) \tag{3.6}$$

[1]From this point forward random variables are no longer underlined

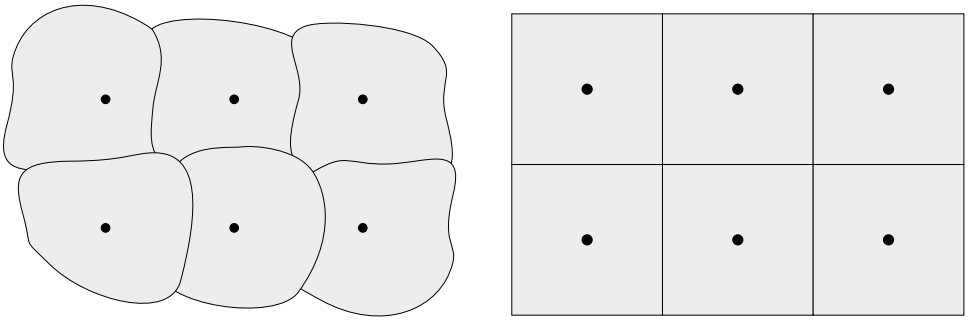

Figure 3.2: Left: Pull-in regions that cover $\mathbb{R}^2$ without gaps and overlap. Right: An example of integer translational invariant pull-in regions that cover $\mathbb{R}^2$ without gaps and overlap.

with the indicator function $s_z(x)$:

$$s_z(x) = \begin{cases} 1 & \text{if } x \in S_z \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

The integer estimator is completely defined by the pull-in region S_z, so that it is possible to define a class of integer estimators by imposing constraints on the pull-in regions. This means that not all integer estimators are allowed, but only estimators that belong to the class of admissible integer estimators. This class is defined as follows.

Definition 3.1.1 *Admissible integer estimators*

An integer estimator, $\breve{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a})$, is said to be admissible when its pull-in region, $S_z = \{x \in \mathbb{R}^n \mid z = S(x)\}$, $z \in \mathbb{Z}^n$, satisfies:

$$(i) \quad \bigcup_{z \in Z^n} S_z = \mathbb{R}^n$$

$$(ii) \quad \text{Int}(S_u) \cap \text{Int}(S_z) = \emptyset, \quad \forall u, z \in \mathbb{Z}^n, u \neq z$$

$$(iii) \quad S_z = z + S_0, \qquad \forall z \in \mathbb{Z}^n$$

where 'Int' denotes the interior of the subset.

The motivation for this definition is illustrated in figure 3.2 and was given in Teunissen (1999a). The first condition states that the collection of all pull-in regions should cover the complete space $\mathbb{R}^n$, so that indeed all real-valued vectors will be mapped to an integer vector. The second condition states that the pull-in regions must be disjunct, i.e. that there should be no overlap between the pull-in regions, so that the float solution is mapped to just one integer vector. Finally, the third condition is that of translational invariance, which means that if the float solution is perturbed by $z \in \mathbb{Z}^n$, the corresponding integer solution is perturbed by the same amount: $S(\hat{a}+z) = S(\hat{a})+z$. This allows one to apply the integer remove-restore technique: $S(\hat{a} - z) + z = S(\hat{a})$.

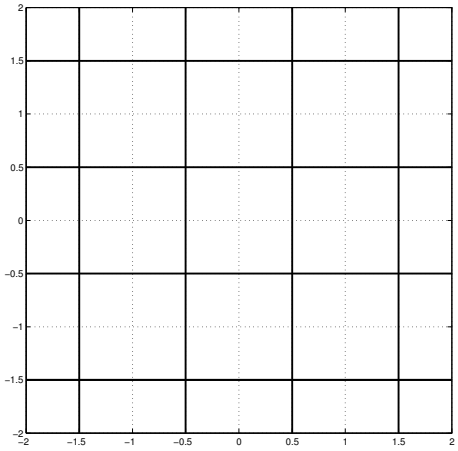

Figure 3.3: Two-dimensional pull-in regions in the case of rounding.

Examples of integer estimators that belong to the class of admissible integer estimators are integer rounding (R), integer bootstrapping (B), and integer least-squares (ILS). Each of these estimators will be described briefly in the following sections.

3.1.1 Integer rounding

The simplest way to obtain an integer vector from the real-valued float solution is to round each entry of $\hat{a}$ to its nearest integer. The corresponding integer estimator, $\check{a}_{\mathsf{R}}$, reads then:

$$\check{a}_{\mathsf{R}} = \begin{pmatrix} [\hat{a}_1] \\ \vdots \\ [\hat{a}_n] \end{pmatrix} \tag{3.8}$$

where $[\cdot]$ denotes rounding to the nearest integer. It can be easily shown that this estimator is admissible, see Teunissen (1999a).

Since each component of the real-valued ambiguity vector is rounded to the nearest integer, the absolute value of the maximum difference between the float and fixed ambiguities is $\frac{1}{2}$. The pull-in region $S_{z,\mathsf{R}}$ that corresponds to this integer estimator is therefore given as:

$$S_{z,\mathsf{R}} = \bigcap_{i=1}^{n} \left\{ x \in \mathbb{R}^n \mid \quad |x_i - z_i| \leq \frac{1}{2} \right\}, \quad \forall z \in \mathbb{Z}^n \tag{3.9}$$

So, the pull-in regions are the n-dimensional unit cubes centered at $z \in \mathbb{Z}^n$. For the two-dimensional case this is visualized in figure 3.3.

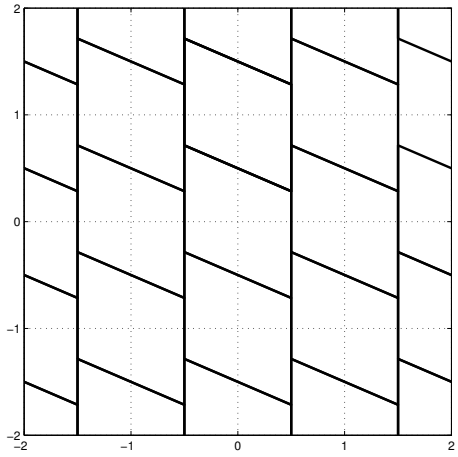

Figure 3.4: Two-dimensional pull-in regions in the case of bootstrapping.

3.1.2 Integer bootstrapping

A generalization of the integer rounding method is the *sequential* integer rounding method, also referred to as the integer bootstrapping method, see e.g (Blewitt 1989; Dong and Bock 1989). In contrast to integer rounding, the integer bootstrapping estimator takes the correlation between the ambiguities into account. It follows from a sequential conditional least-squares adjustment with a conditioning on the integer ambiguity values from the previous steps (Teunissen 1993; Teunissen 1998b). The components of the bootstrapped estimator are given as:

$$\begin{aligned}
\breve{a}_{1,\mathrm{B}} &= [\hat{a}_1] \\
\breve{a}_{2,\mathrm{B}} &= [\hat{a}_{2|1}] = [\hat{a}_2 - \sigma_{\hat{a}_2\hat{a}_1}\sigma_{\hat{a}_1}^{-2}(\hat{a}_1 - \breve{a}_{1,\mathrm{B}})] \\
&\vdots \\
\breve{a}_{n,\mathrm{B}} &= [\hat{a}_{n|N}] = \left[\hat{a}_n - \sum_{i=1}^{n-1}\sigma_{\hat{a}_n\hat{a}_{i|I}}\sigma_{\hat{a}_{i|I}}^{-2}(\hat{a}_{i|I} - \breve{a}_{i,\mathrm{B}})\right]
\end{aligned} \tag{3.10}$$

where $\hat{a}_{i|I}$ stands for the ith ambiguity obtained through a conditioning on the previous $I = \{1, \ldots, (i-1)\}$ sequentially rounded ambiguities. Since, the first entry is simply rounded to the nearest integer, one should start with the most precise float ambiguity. The real-valued sequential least-squares solution can be obtained by means of the triangular decomposition of the vc-matrix of the ambiguities: $Q_{\hat{a}} = LDL^T$, where L denotes a unit lower triangular matrix with entries

$$l_{j,i} = \sigma_{\hat{a}_j\hat{a}_{i|I}}\sigma_{\hat{a}_{i|I}}^{-2} \tag{3.11}$$

and D a diagonal matrix with the conditional variances $\sigma_{\hat{a}_{i|I}}^2$ as its entries. Since by definition $|\hat{a}_{i|I} - [\hat{a}_{i|I}]| \leq \frac{1}{2}$, the pull-in region $S_{z,\mathrm{B}}$ that belongs to the bootstrapped

estimator is given by:

$$S_{z,\mathrm{B}} = \bigcap_{i=1}^{n} \left\{ x \in \mathbb{R}^n \mid |c_i^T L^{-1}(x-z)| \leq \frac{1}{2},\ \forall z \in \mathbb{Z}^n \right\} \tag{3.12}$$

with c_i the canonical vector having a one as its ith entry. Note that the bootstrapped and rounded solution become identical if the ambiguity vc-matrix, $Q_{\hat{a}}$, is diagonal.

The shape of the pull-in region in the two-dimensional case is a parallelogram, see figure 3.4. For more dimensions it will be the multivariate version of a parallelogram.

3.1.3 Integer least-squares

Optimizing on the integer nature of the ambiguity parameters, cf.(Teunissen 1999a), involves solving a non-standard least-squares problem, referred to as *integer least-squares* (ILS) in Teunissen (1993). The solution is obtained by solving the following minimization problem:

$$\min_{z,\zeta} \|y - Az - B\zeta\|^2_{Q_y}, \quad z \in \mathbb{Z}^n,\ \zeta \in \mathbb{R}^p \tag{3.13}$$

The following orthogonal decomposition can be used:

$$\|y - Az - B\zeta\|^2_{Q_y} = \underbrace{\|\hat{e}\|^2_{Q_y}}_{(3.2)} + \underbrace{\|\hat{a} - z\|^2_{Q_{\hat{a}}}}_{(3.3)} + \underbrace{\|\hat{b}(z) - \zeta\|^2_{Q_{\hat{b}|\hat{a}}}}_{(3.4)} \tag{3.14}$$

with the residuals of the float solution $\hat{e} = y - A\hat{a} - B\hat{b}$, and the conditional baseline estimator $\hat{b}(z) = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}}^{-1}(\hat{a} - z)$.

It follows from equation (3.14) that the solution of the minimization problem in equation (3.13) is solved using the three step procedure described in section 3.1. The first term on the right-hand side follows from the float solution (3.2). Taking into account the integer nature of the ambiguities means that the second term on the right-hand side of equation (3.14) needs to be minimized, conditioned on $z \in \mathbb{Z}^n$. This corresponds to the ambiguity resolution step in (3.3), so that $z = \check{a}$. Finally, solving for the last term corresponds to fixing the baseline as in equation (3.4). Then the last term in equation (3.14) becomes equal to zero, so that indeed the minimization problem is solved.

The integer least-squares problem focuses on the second step and can now be defined as:

$$\check{a}_{\mathrm{LS}} = \arg\min_{z \in \mathbb{Z}^n} \|\hat{a} - z\|^2_{Q_{\hat{a}}} \tag{3.15}$$

where $\check{a}_{\mathrm{LS}} \in \mathbb{Z}^n$ is the fixed ILS ambiguity solution. This solution cannot be 'computed' as with integer rounding and integer bootstrapping. Instead an integer search is required to obtain the solution.

The ILS pull-in region $S_{z,\mathrm{LS}}$ is defined as the collection of all $x \in \mathbb{R}^n$ that are closer to z than to any other integer grid point in $\mathbb{R}^n$, where the distance is measured in the

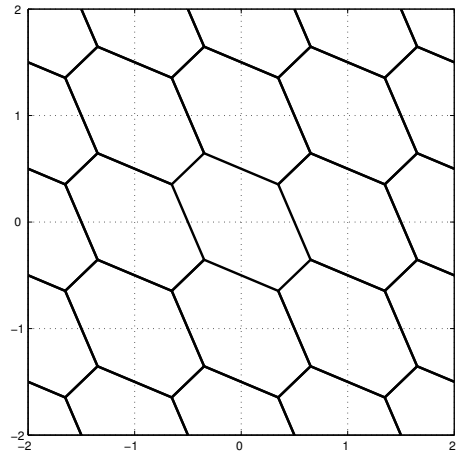

Figure 3.5: Two-dimensional pull-in regions in the case of integer least-squares.

metric of the vc-matrix $Q_{\hat{a}}$. The pull-in region that belongs to the integer z follows as:

$$S_{z,\mathsf{LS}} = \left\{ x \in \mathbb{R}^n \mid \|x - z\|^2_{Q_{\hat{a}}} \leq \|x - u\|^2_{Q_{\hat{a}}}, \forall u \in \mathbb{Z}^n \right\} \tag{3.16}$$

A similar representation as for the bootstrapped pull-in region can be obtained by using:

$$\|x - z\|^2_{Q_{\hat{a}}} \leq \|x - u\|^2_{Q_{\hat{a}}} \Longleftrightarrow (u - z)^T Q^{-1}_{\hat{a}}(x - z) \leq \frac{1}{2}\|u - z\|^2_{Q_{\hat{a}}}, \forall u \in \mathbb{Z}^n$$

With this result, it follows that

$$S_{z,\mathsf{LS}} = \bigcap_{c \in \mathbb{Z}^n} \left\{ x \in \mathbb{R}^n \mid |c^T Q^{-1}_{\hat{a}}(x - z)| \leq \frac{1}{2}\|c\|^2_{Q_{\hat{a}}} \right\} \forall z \in \mathbb{Z}^n \tag{3.17}$$

The ILS pull-in regions are thus constructed as intersecting half-spaces, which are bounded by the planes orthogonal to $(u - z)$, $u \in \mathbb{Z}^n$ and passing through $\frac{1}{2}(u + z)$. It can be shown that at most $2^n - 1$ pairs of such half-spaces are needed for the construction. The integer vectors u must then be the $2^n - 1$ adjacent integers. So, for the 2-dimensional case this means three pairs are needed, and the ILS pull-in regions are hexagons, see figure 3.5.

3.1.4 The LAMBDA method

The ILS procedure is mechanized in the LAMBDA (Least-Squares AMBiguity Decorrelation Adjustment) method, see (Teunissen 1993; Teunissen 1995; De Jonge and Tiberius 1996).

As mentioned in section 3.1.3, the integer ambiguity solution is obtained by an integer search. Therefore, a search space is defined as:

$$\Omega_a = \left\{ a \in \mathbb{Z}^n \mid (\hat{a} - a)^T Q^{-1}_{\hat{a}}(\hat{a} - a) \leq \chi^2 \right\} \tag{3.18}$$

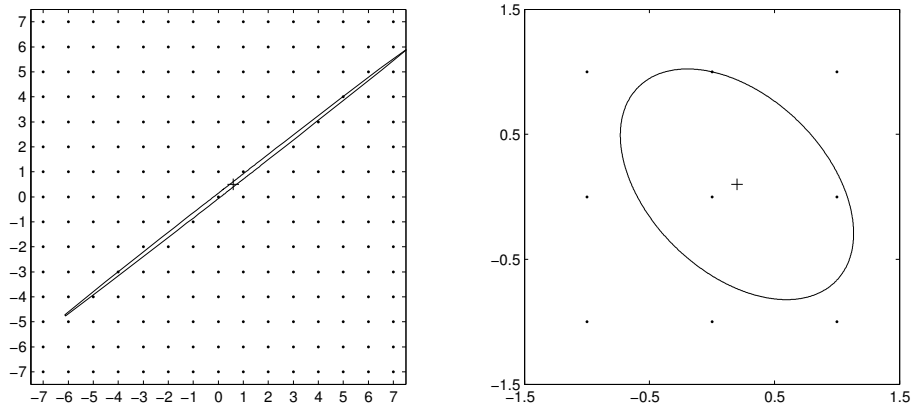

Figure 3.6: Two-dimensional example of the search space before (left) and after (right) the decorrelating Z-transformation. The float ambiguity vector is depicted with a $+$-sign.

with χ^2 a positive constant that needs to be chosen. The search space is an n-dimensional ellipsoid centered at $\hat{a}$. Its shape is governed by the vc-matrix $Q_{\hat{a}}$ and its size by χ^2. Due to the high correlation between the individual ambiguities, the search space in the case of GNSS is extremely elongated, so that the search for the integer solution may take very long. Therefore, the search space is first transformed to a more spherical shape by means of a decorrelation of the original float ambiguities. This decorrelation is attained by a transformation:

$$\hat{z} = Z^T \hat{a}, \quad Q_{\hat{z}} = Z^T Q_{\hat{a}} Z \quad \text{and} \quad \check{a} = Z^{-T} \check{z} \tag{3.19}$$

This transformation needs to be admissible,i.e. Z and Z^{-1} must have integer entries, so that the integer nature of the ambiguities is preserved and the Z-transformation is volume-preserving with respect to the search space, since the determinant of Z is equal to ± 1. The transformed search space is then given by:

$$\Omega_z = \left\{ z \in \mathbb{Z}^n \mid (\hat{z} - z)^T Q_{\hat{z}}^{-1} (\hat{z} - z) \leq \chi^2 \right\} \tag{3.20}$$

Figure 3.6 shows a two-dimensional example of the search space before and after the decorrelating Z-transformation. Obviously, the search tree that is required to find the three integers in the search space will be much smaller after decorrelation.

Using the LDL^T-decomposition of $Q_{\hat{z}}$, the left-hand side of equation (3.20) can be written as a sum-of-squares:

$$\sum_{i=1}^{n} \frac{(\hat{z}_{i|I} - z_i)^2}{\sigma_{i|I}^2} \leq \chi^2 \tag{3.21}$$

with the conditional LS-estimator $\hat{z}_{i|I}$, and the conditional variances $\sigma_{i|I}^2$ that come from

matrix D. From equation (3.21) the n intervals that are used for the search follow as:

$$\begin{aligned}(\hat{z}_1 - z_1)^2 &\leq \sigma_1^2 \chi^2 \\ (\hat{z}_{2|1} - z_2)^2 &\leq \sigma_{2|1}^2 (\chi^2 - \frac{(\hat{z}_1 - z_1)^2}{\sigma_1^2}) \\ &\vdots\end{aligned} \tag{3.22}$$

Besides the shape of the search space, also the size is important in order for the search to be efficient. That means that it should not contain too many integer grid points. The choice of χ^2 should therefore be such that the search space is small but at the same time such that it is guaranteed that it contains at least one integer, or two integers in the case the solution will be validated, see section 3.5. Since the bootstrapped solution gives a good approximation of the ILS estimator and is easy to compute, and by definition $\|\hat{z} - \check{z}_{\mathsf{LS}}\|_{Q_{\hat{z}}}^2 \leq \|\hat{z} - \check{z}_{\mathsf{B}}\|_{Q_{\hat{z}}}^2$, the size of the search space can be chosen as:

$$\chi^2 = (\hat{z} - \check{z}_{\mathsf{B}})^T Q_{\hat{z}}^{-1} (\hat{z} - \check{z}_{\mathsf{B}}) \tag{3.23}$$

If more than one candidate must be determined, bootstrapping can still be used to set χ^2. The conditional estimates $\hat{z}_{i|I}$ are then not only rounded to their nearest integer, but also to the second nearest integer. In this way $n+1$ conditionally rounded integers $\check{z}_{\mathsf{B},j}$ are obtained, for which also the squared norms $R_j = \|\hat{z} - \check{z}_{\mathsf{B},j}\|_{Q_{\hat{z}}}^2$ are determined. They are collected in a vector $(R_1 \cdots R_{n+1})$ with the ordering $R_i \leq R_{i+1}$, $\forall i = 1, \ldots, n$. If the requested number of candidates is $l \leq n+1$, the size of the search space is chosen as $\chi^2 = R_l$ since it is then guaranteed that the search space contains at least l candidates.

Note that this approach to set the size of the search space is different from the one described in De Jonge and Tiberius (1996), which is based on the relation between the volume of the ellipsoid and the number of candidates it contains. This may result in much larger search spaces. Note, however, that if more than $n+1$ candidates are requested, the approximation based on the volume of the search space must be used. In Teunissen et al. (1996) it was shown that the volume, V_n, is a good indicator of the number of integers, N_z, contained in an ellipsoidal region if this number is larger than a few:

$$N_z \approx [V_n] \tag{3.24}$$

with

$$V_n = \lambda^n U_n \sqrt{|Q_{\hat{z}}|} \tag{3.25}$$

and U_n the volume of the unit sphere in $\mathbb{R}^n$ which is given as $U_n = \pi^{\frac{n}{2}} / \Gamma(\frac{n}{2}+1)$, where $\Gamma(x)$ is the gamma function, see equation (A.11).

Once the size of the search space is set, the actual search can be carried out. In De Jonge and Tiberius (1996), it is described how to do that in an efficient way.

Table 3.1: Overview of ambiguity resolution methods.

Method	Name	References
Least-Squares Ambiguity Search Technique	LSAST	Hatch (1990)
Fast Ambiguity Resolution Approach	FARA	Frei and Beutler (1990)
Modified Cholesky decomposition		Euler and Landau (1992)
Least-squares AMBiguity Decorrelation Adjustment	LAMBDA	Teunissen (1993)
Null method		Martín-Neira et al. (1995); Fernàndez-Plazaola et al. (2004)
Fast Ambiguity Search Filter	FASF	Chen and Lachapelle (1995)
Three Carrier Ambiguity Resolution	TCAR	Harris (1997)
Integrated TCAR		Vollath et al. (1998)
Optimal Method for Estimating GPS Ambiguities	OMEGA	Kim and Langley (1999)
Cascade Integer Resolution	CIR	Jung et al. (2000)

3.1.5 Other ambiguity resolution methods

Besides LAMBDA, several other ambiguity resolution methods have been described in literature. Table 3.1.5 gives an overview of some well-known methods with references. Only TCAR and CIR are based on the bootstrapping estimator, all other methods are based on the ILS principle of minimizing the squared norm of residuals. The methods essentially differ in the way the search space is defined. In Kim and Langley (2000), some of the methods were conceptually compared. A comparison of LAMBDA with CIR, TCAR, ITCAR and the Null-method is made in Joosten and Verhagen (2003), Verhagen and Joosten (2004).

3.2 Quality of the integer ambiguity solution

The uncertainty of parameter estimators is captured by their parameter distribution. For normally distributed data, the uncertainty is completely captured by the vc-matrix. However, the GNSS model contains integer parameters to which this does not apply. In this section, the distributional properties of both the real-valued and the integer ambiguity parameters are described.

An important measure of the reliability of the fixed solution is the probability of correct integer estimation, the success rate. Once the parameter distribution of the integer estimator is given, the success rate can be determined. This will be the subject of section 3.2.2.

3.2.1 Parameter distributions of the ambiguity estimators

The float ambiguities are assumed to be normally distributed, i.e. $\hat{a} \sim N(a, Q_{\hat{a}})$, see appendix A.2. The probability density function (PDF) of $\hat{a}$ reads then

$$f_{\hat{a}}(x) = \frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \exp\{-\frac{1}{2}\|x-a\|^2_{Q_{\hat{a}}}\} \tag{3.26}$$

where $|\cdot|$ denotes the determinant. The joint PDF of the float and fixed ambiguities is given by:

$$f_{\hat{a},\check{a}}(x,z) = f_{\hat{a}}(x)s_z(x), \quad x \in \mathbb{R}^n, z \in \mathbb{Z}^n \tag{3.27}$$

with $s_z(x)$ the indicator function of equation (3.7). For the proof see (Teunissen 2002). The marginal distributions of $\hat{a}$ and $\check{a}$ can be recovered from this joint distribution by integration or summation over respectively x and z. This gives the following distribution function of the fixed ambiguities:

$$P(\check{a}=z) = \int_{\mathbb{R}^n} f_{\hat{a},\check{a}}(x,z)dx = \int_{S_z} f_{\hat{a}}(x)dx, \quad z \in \mathbb{Z}^n \tag{3.28}$$

It is a probability mass function (PMF), and is referred to as the integer normal distribution for the ILS estimator. The three top panels of figure 3.7 show the marginal and joint distributions of $\hat{a}$ and $\check{a}$.

The distribution of $\check{a}$ has some (desirable) properties, cf.(Teunissen 1998a; Teunissen 1998c). The first is that for all admissible integer estimators the distribution is symmetric about a:

$$P(\check{a}=a-z) = P(\check{a}=a+z), \quad \forall z \in \mathbb{Z}^n \tag{3.29}$$

From this property it follows that all admissible estimators are unbiased:

$$E\{\check{a}\} = \sum_{z\in\mathbb{Z}^n} zP(\check{a}=z) = a \tag{3.30}$$

Another important property is that if a is the correct integer, the probability that this integer is obtained as solution is always larger than the probability that another integer is estimated:

$$P(\check{a}=a) > P(\check{a}=z,\, z \in \mathbb{Z}^n \backslash \{a\}) \tag{3.31}$$

Finally, the vc-matrix of the fixed ambiguities follows from equation (3.28) as:

$$D\{\check{a}\} = \sum_{z\in\mathbb{Z}^n} (z-a)(z-a)^T P(\check{a}=z) = Q_{\check{a}} \tag{3.32}$$

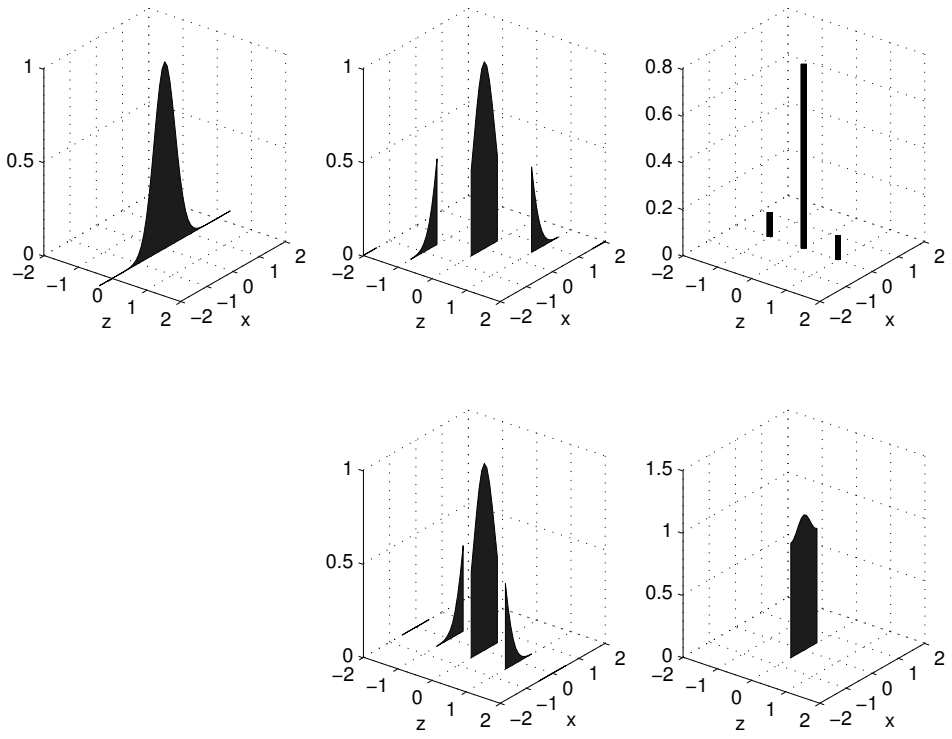

Figure 3.7: The joint and marginal distribution of $\hat{a}$, $\check{a}$ and $\check{\epsilon}$: PDF $f_{\hat{a}}(x)$ (top left); joint PDF $f_{\hat{a},\check{a}}(x,z)$ (top middle); PMF $P(\check{a}=z)$ (top right); PDF $f_{\check{\epsilon},\check{a}}(x,z)$ (bottom middle); PDF $f_{\check{\epsilon}}(x)$ (bottom right).

3.2.2 Success rates

The success rate, P_s, equals the probability that the ambiguities are fixed to the correct integers. It follows from equation (3.28) as:

$$P_s = P(\check{a}=a) = \int_{S_a} f_{\hat{a}}(x)dx \tag{3.33}$$

It is a very important measure, since the fixed solution should only be used if there is enough confidence in this solution. In Teunissen (1999a) it was proven that:

$$P(\check{a}_{\mathsf{LS}}=a) \geq P(\check{a}=a) \tag{3.34}$$

for any admissible integer estimator $\check{a}$. Therefore, the ILS estimator is optimal in the class of admissible integer estimators. Furthermore, it was shown in Teunissen (1998d) that:

$$P(\check{a}_{\mathsf{R}}=a) \leq P(\check{a}_{\mathsf{B}}=a) \tag{3.35}$$

Unfortunately, it is very complicated to evaluate equation (3.33) for the integer least-squares estimator because of the complex integration region. Therefore, approximations

have to be used. The most important approximations will be outlined here. An evaluation is given at the end of this subsection.

Lower bound based on bootstrapping

The bootstrapped estimator has a close to optimal performance after decorrelation of the ambiguities using the Z-transformation of the LAMBDA method. Therefore, the bootstrapped success rate can be used as lower bound for the ILS success rate as was proposed in Teunissen (1999a), since it is possible to give an exact evaluation of this success rate, (Teunissen 1998d):

$$\begin{aligned} P_{s,\mathsf{B}} &= P(\bigcap_{i=1}^{n} |\hat{a}_{i|I} - a_i| \leq \frac{1}{2}) \\ &= \prod_{i=1}^{n} P([\hat{a}_{i|I}] = a_i \;\Big|\; [\hat{a}_1] = a_1, \ldots, [\hat{a}_{i-1|I-1}] = a_{i-1}) \\ &= \prod_{i=1}^{n} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{\sigma_{i|I}\sqrt{2\pi}} \exp\{-\frac{1}{2}\left(\frac{x}{\sigma_{i|I}}\right)^2\} dx \\ &= \prod_{i=1}^{n} \left(2\Phi(\frac{1}{2\sigma_{i|I}}) - 1\right) \end{aligned} \tag{3.36}$$

with $\Phi(x)$ the cumulative normal distribution of equation (A.7), and $\sigma_{i|I}$ the standard deviation of the ith least-squares ambiguity obtained through a conditioning on the previous $I = \{1, ..., i-1\}$ ambiguities. These conditional standard deviations are equal to the square root of the entries of matrix D from the LDL^T-decomposition of the vc-matrix $Q_{\hat{a}}$. In order to obtain a strict lower bound of the ILS success rate, the bootstrapped success rate in equation (3.37) must be computed for the decorrelated ambiguities $\hat{z}$ and corresponding conditional standard deviations. Hence:

$$P_{s,\mathsf{LS}} \geq \prod_{i=1}^{n} \left(2\Phi(\frac{1}{2\sigma_{i|I}}) - 1\right) \tag{3.37}$$

with $\sigma_{i|I}$ the square roots of the diagonal elements of D from $Q_{\hat{z}} = LDL^T$.

Upper bound based on ADOP

The Ambiguity Dilution of Precision (ADOP) is defined as a diagnostic that tries to capture the main characteristics of the ambiguity precision. It is defined as:

$$ADOP = \sqrt{|Q_{\hat{a}}|}^{\frac{1}{n}} = \prod_{i=1}^{n} \sigma_{i|I}^{\frac{1}{n}} \tag{3.38}$$

and has units of $cycles$. It was introduced in Teunissen (1997), and described and analyzed in Teunissen and Odijk (1997). The ADOP measure has some desirable properties. First, it is invariant within the class of admissible ambiguity transformations, which means for example that ADOP is independent of the chosen reference satellite in the double difference approach, and ADOP will not change after the decorrelating Z-transformation of the ambiguities. When the ambiguities are completely decorrelated, the ADOP equals the geometric mean of the standard deviations of the ambiguities, hence ADOP approximates the average precision of the transformed ambiguities.

In Teunissen (2000b) it was proven that an upper bound for the ILS success rate can be given based on the ADOP:

$$P_{s,\mathrm{LS}} \leq P\left(\chi^2(n,0) \leq \frac{c_n}{ADOP^2}\right) \tag{3.39}$$

with

$$c_n = \frac{\left(\frac{n}{2}\Gamma(\frac{n}{2})\right)^{\frac{2}{n}}}{\pi}$$

This upper bound is identical to the one presented without proof in Hassibi and Boyd (1998).

Lower and upper bounds based on bounding the integration region

In Teunissen (1998c) lower and upper bounds for the ILS success rate were obtained by bounding the integration region. Obviously, a lower bound is obtained if the integration region is chosen such that it is completely contained by the pull-in region, and an upper bound is obtained if the integration region is chosen such that it completely contains the pull-in region. The integration region can then be chosen such that the integral is easy-to-evaluate. In (ibid) the integration region for the lower bound was chosen as an ellipsoidal region $C_a \subset S_a$.

The upper bound can be obtained by defining a region $U_a \supset S_a$, with S_a as defined in (3.17):

$$S_a = \{\hat{a} \in \mathbb{R}^n \mid |w| \leq \frac{1}{2}\|c\|_{Q_{\hat{a}}}, \forall c \in \mathbb{Z}^n\} \quad \text{with} \quad w = \frac{c^T Q_{\hat{a}}^{-1}(\hat{a}-a)}{\|c\|_{Q_{\hat{a}}}} \tag{3.40}$$

Note that the w in this expression can be geometrically interpreted as the orthogonal projection of $(\hat{a}-a)$ onto the direction vector c. Hence, S_a is the intersection of *banded* subsets centered at a and having a width $\|c\|_{Q_{\hat{a}}}$. Any *finite* intersection of these banded subsets encloses S_a, and therefore the subset U_a could be chosen as

$$U_a = \{\hat{a} \in \mathbb{R}^n \mid |w_i| \leq \frac{1}{2}\|c_i\|_{Q_{\hat{a}}},\, i = 1, \ldots, p\} \supset S_a$$

with $w_i \sim N(0,1)$. The choice for p is still open, but a larger p will result in a sharper upper bound for the success rate. However, if $p > 1$ the w_i are correlated. This is handled by defining a p-vector v as

$$v = (v_1, \ldots, v_p)^T \quad \text{with} \quad v_i = \frac{w_i}{\|c_i\|_{Q_{\hat{a}}}}$$

Then $U_a = \{\hat{a} \in \mathbb{R}^n \mid \bigcap_{i=1}^{p} |v_i| \leq \frac{1}{2}\}$. Therefore, the probability $P(\hat{a} \in U_a)$ equals the probability that component-wise rounding of the vector v produces the zero vector. This means that $P(\hat{a} \in U_a)$ is bounded from above by the probability that conditional rounding, cf.(Teunissen 1998d), produces the zero vector, i.e.:

$$P_{s,\text{LS}} \leq P(\hat{a} \in U_a) \leq \prod_{i=1}^{p} \left(2\Phi(\frac{1}{2\sigma_{v_{i|I}}}) - 1 \right) \tag{3.41}$$

with $\sigma_{v_{i|I}}$ the conditional standard deviation of v_i. The conditional standard deviations are equal to the diagonal entries of the matrix D from the LDL^T-decomposition of the vc-matrix of v. The elements of this vc-matrix are given as:

$$\sigma_{v_i v_j} = \frac{c_i^T Q_{\hat{a}}^{-1} c_j}{\|c_i\|_{Q_{\hat{a}}}^2 \|c_j\|_{Q_{\hat{a}}}^2}$$

In order to avoid the conditional standard deviations becoming zero, the vc-matrix of v must be of full rank, and thus the vectors c_i, $i = 1, \ldots, p \leq n$ need to be linearly independent.

The procedure for the computation of this upper bound is as follows. LAMBDA is used to find the $q >> n$ closest integers $c_i \in \mathbb{Z}^n \setminus \{0\}$ for $\hat{a} = 0$. These q integer vectors are ordered by increasing distance to the zero vector, measured in the metric $Q_{\hat{a}}$. Start with $C = c_1$, so that $rank(C) = 1$. Then find the first candidate c_j for which $rank(c_1\ c_j) = 2$. Continue with $C = (c_1\ c_j)$ and find the next candidate that results in an increase in rank. Continue this process until $rank(C) = n$.

In Kondo (2003) the correlation between the w_i is not taken into account, which means that instead of the conditional variances, simply the variances of the v_i are used. Then the following is obtained:

$$\prod_{i=1}^{p} \left(2\Phi(\frac{1}{2\sigma_{v_i}}) - 1 \right) = \prod_{i=1}^{p} P_{s,i}$$

with the $P_{s,i}$ equal to

$$P_{s,i} = \frac{2}{\sqrt{2\pi}\sigma_{v_i}} \int_0^{\frac{1}{2}} \exp\left\{ -\frac{1}{2} \frac{x^2}{\sigma_{v_i}^2} \right\} dx \tag{3.42}$$

It is known, cf.(Teunissen 1998d), that

$$\prod_{i=1}^{p} \left(2\Phi(\frac{1}{2\sigma_{v_i}}) - 1 \right) \leq P(\hat{a} \in U_a)$$

This means that it is only guaranteed that Kondo's approximation of the success rate is a lower bound if $P(\hat{a} \in U_a)$ is equal to the success rate. This will be the case if p is

chosen equal to half the number of facets that bound the ILS pull-in region. So, it is required to know this number, but in practice only the bounds are known:

$$n \leq p \leq 2^n - 1$$

If p is chosen to be smaller than half the number of bounding facets, it is not guaranteed that the approximation gives a lower bound. On the other hand, if p is chosen to be larger than required in order to guarantee that $P(\hat{a} \in U_a) = P(\hat{a} \in S_a)$, the lower bound is less strict since the lower bound is defined as a product of probabilities which are all smaller or equal to one. Note that p may become very large when many satellites are visible. For instance, with 6 visible satellites and two frequencies available, the number of unknown ambiguities for one epoch is $n = 10$, and $p \leq 2^n - 1 = 1023$.

It is possible to find *all* adjacent integers, but it is computationally demanding. First, note that it is not always the case that the $2p$ adjacent integers are also the $2p$ closest integers. Therefore, a large set of integers c_i must be selected, in the same way as for the computation of the upper bound described above, but now with $q >> 2(2^n - 1)$. For each integer in this set it must be checked if it is adjacent, which is the case if $\frac{1}{2}c_i$ lies on the boundary of both the pull-in regions S_0 and S_{c_i}. This is the case if:

$$\|\frac{1}{2}c_i - 0\|^2_{Q_{\hat{a}}} = \|\frac{1}{2}c_i - c_i\|^2_{Q_{\hat{a}}} = \min_{z \in \mathbb{Z}^n} \|\frac{1}{2}c_i - z\|^2_{Q_{\hat{a}}} \tag{3.43}$$

Note that if c_j is selected as adjacent integer, $-c_j$ must not be included in the set $C = (c_1 \ \ldots, c_p)$ adjacent integers that is used to compute the lower bound.

Approximation based on ADOP

Until now only lower and upper bounds of the ILS success rate were given. If these bounds are sharp, it is very useful to know whether an approximation is a lower or upper bound, since in general a user will choose a minimum required success rate. If the lower bound is below this minimum, the user will not have enough confidence in the fixed solution – although there is still a probability that the actual success rate is large enough. An upper bound could be used in a similar way: if it is below a certain threshold value, one cannot expect ambiguity resolution to be successful.

Instead of a lower and/or upper bound, it is also possible to use the following approximation:

$$P_{s,\mathrm{LS}} \approx (2\Phi(\frac{1}{2ADOP}) - 1)^n \tag{3.44}$$

In Teunissen (2003b) it was proven that this is an invariant upper bound of the bootstrapped success rate, and it is equal to the bootstrapped success rate when all ambiguities are completely decorrelated. Since the bootstrapped success rate is used as a lower bound for the ILS success rate, the bootstrapped upper bound is expected to be a good approximation of the ILS success rate. The reason is that the ADOP approximates the average precision of the transformed ambiguities, and can therefore be used to replace the individual conditional standard deviations in equation (3.37).

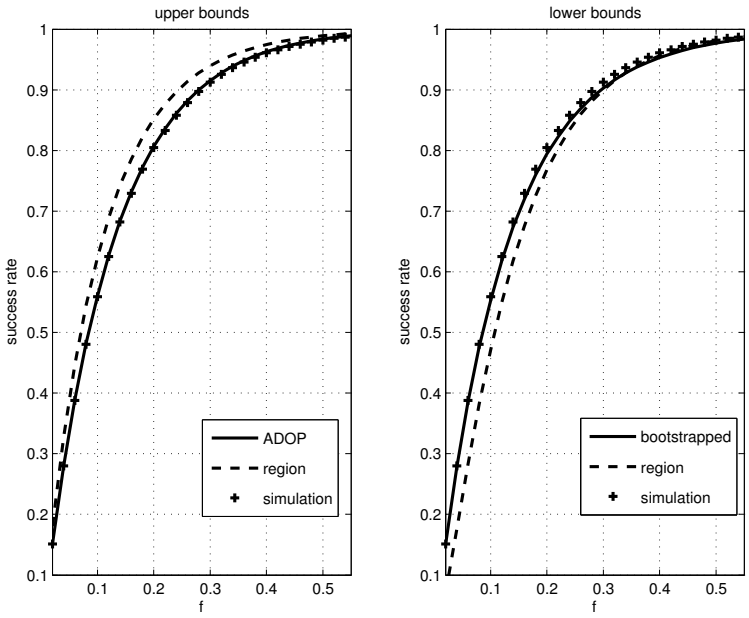

Figure 3.8: Upper and lower bounds for the success rate in the 2-D case as function of f with vc-matrix $\frac{1}{f}Q_{\hat{a},\text{ref}}$.

Evaluation of the approximations

In order to evaluate the lower and upper bounds of the success rate, simulations are used. The procedure is explained in appendix B. The results were also presented in Verhagen (2003).

First the two-dimensional case is considered, using:

$$Q_{\hat{z}} = \frac{1}{f}\begin{bmatrix} 0.0216 & -0.0091 \\ -0.0091 & 0.0212 \end{bmatrix}, \quad 0 < f \leq 1$$

for different values of f. The results are shown in figure 3.8. The left panel shows the two upper bounds and the success rates from the simulations. Obviously, the ADOP-based upper bound is very strict and is always much better than the upper bound based on bounding the integration region. The panels on the right show the lower bounds. It follows that for lower success rates (< 0.93) the bootstrapped success rate is the best lower bound. For higher success rates the success rate proposed by Kondo works very well and is better than the bootstrapped lower bound.

Table 3.2 shows the maximum and mean differences of the approximations with the success rate from simulation. From these differences it follows that the ADOP-based approximation, the bootstrapped lower bound and the ADOP-based upper bound are best.

Because of its simplicity the geometry-free model is very suitable for a first evaluation,

Table 3.2: Two-dimensional example. Mean and maximum difference between success rate, P_s, based on simulations and the approximations. The success rate for which the maximum difference is obtained is given in the last row.

	LB bootstr.	LB region	UB ADOP	UB region	ADOP
mean difference	0.005	0.018	0.001	0.018	0.004
max. difference	0.010	0.105	0.003	0.065	0.009
P_s	0.805	0.388	0.833	0.559	0.805

Table 3.3: Approximated success rates using simulations (sim.), the lower bounds based on bootstrapping (LB bootstr.) and bounding the integration region (LB region), the upper bounds based on ADOP (UB ADOP) and bounding the integration region (UB region), and the approximation based on ADOP (ADOP).

example	sim.	LB bootstr.	LB region	UB ADOP	UB region	ADOP
2-D	1.000	0.999	1.000	1.000	1.000	0.999
06_01	0.818	0.749	0.698	0.848	0.942	0.765
06_02	0.442	0.410	0.118	0.475	0.626	0.417
10_01	0.989	0.976	0.988	0.999	0.992	0.978
10_03	0.476	0.442	0.126	0.681	0.661	0.525

though it is more useful to know how well the approximations work in practice. Therefore, simulations were carried out for several geometry-based models, see appendix B. The resulting lower and upper bounds are shown in table 3.3. The first row shows the results for the two-dimensional vc-matrix with $f = 1$.

The results show that Kondo's lower bound works very well for a high success rate, but in general the bootstrapped lower bound is much better. It can be concluded that Kondo's lower bound seems to be useful only in a few cases. Firstly, to obtain a strict lower bound the precision should be high, so that the success rate is high. Even then, it depends on the minimum required success rate whether or not it is really necessary to use the approximation: if the bootstrapped success rate is somewhat lower than this minimum required success rate, Kondo's approximation can be used to see if it is larger. The minimum required success rate could be chosen such that the fixed ambiguity estimator can be considered deterministic. In this case, the discrimination tests as used in practice can be used, see section 3.5.

An advantage of the bootstrapped success rate is that it is very easy to compute, since the conditional variances are already available when using the LAMBDA method. The computation of Kondo's lower bound is more complex, since for high-dimensional problems the number of facets that bound the pull-in region can be very large.

It is difficult to say which upper bound is best. For the examples with only four visible satellites the ADOP-based upper bound is better than the one obtained by bounding the integration region, but in the examples with more satellites the latter is somewhat better. All bounds are best in the case of high precisions, i.e. high success rates. All in all, one can have a little more confidence in the ADOP-based bound, since its overall performance, based on all examples, is slightly better. An advantage of the ADOP-based upper bound is that it is easy to compute, whereas using the upper bound based

on bounding the integration region brings the problem of determining the n closest independent integers to the zero vector.

The approximation based on ADOP performs quite well for the examples shown here. Except for example 10_03, this approximation is smaller than the empirical success rate, and better than the lower bound. The vc-matrix of example 10_03 is obtained for a relatively long baseline and its dimension is equal to 10. This results in conditional standard deviations of the transformed ambiguities that may differ considerably from the approximate average given by ADOP. This explains why the approximation is not so good in this case.

3.2.3 Bias-affected success rates

In section 2.5.1 the minimal detectable effect (MDE) was introduced as a measure of the impact of a certain model error on the final solution. Using equation (2.77) the impact on the float solution can be computed, but for high-precision GNSS applications it would be more interesting to know the impact on the fixed solution. For that purpose, one could compute the effect of a bias in the float ambiguity solution, $\nabla\hat{a}$, on the success rate. This so-called *bias-affected* success rate can be computed once $\nabla\hat{a}$ is known, see (Teunissen 2001a). The bootstrapped success rate in the presence of biases, and after decorrelation so that the bias is $\nabla\hat{z} = Z^T\nabla\hat{a}$, becomes:

$$\begin{aligned} P_\nabla(\check{z}_\mathrm{B} = z) &= \int\limits_{S_{0,\mathrm{B}}} \frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \exp\{-\frac{1}{2}\|x - \nabla\hat{z}\|^2_{Q_{\hat{z}}}\}dx \\ &= \int\limits_{F^{-1}(S_{0,\mathrm{B}})} \frac{1}{\sqrt{|D|}(2\pi)^{\frac{1}{2}n}} \exp\{-\frac{1}{2}\|y - L^{-1}\nabla\hat{z}\|^2_{D}\}dy, \end{aligned} \tag{3.45}$$

where the transformation $F : y = Lx$ was used, so that the bootstrapped pull-in region is transformed as:

$$F^{-1}(S_{0,\mathrm{B}}) = \{y \in \mathbb{R}^n \mid |c_i^T y| \leq \frac{1}{2},\ i = 1, \ldots, n\}$$

with c_i the vector with a one as ith entry and zeros otherwise. The transformed pull-in region equals an origin-centered cube with all sides equal to 1. Since D is a diagonal matrix, the multivariate integral can be written as:

$$\begin{aligned} P_\nabla(\check{z}_\mathrm{B} = z) &= \prod_{i=1}^{n} \int\limits_{|y_i| \leq \frac{1}{2}} \frac{1}{\sigma_{i|I}\sqrt{2\pi}} \exp\{-\frac{1}{2}\left(\frac{y_i - c_i^T L^{-1}\nabla\hat{z}}{\sigma_{i|I}}\right)^2\}dy_i \\ &= \prod_{i=1}^{n} \int\limits_{-\frac{1+2c_i^T L^{-1}\nabla\hat{z}}{2\sigma_{i|I}}}^{\frac{1-2c_i^T L^{-1}\nabla\hat{z}}{2\sigma_{i|I}}} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}v^2\}dv \\ &= \prod_{i=1}^{n} \left(\Phi\left(\frac{1 - 2c_i^T L^{-1}\nabla\hat{z}}{2\sigma_{i|I}}\right) + \Phi\left(\frac{1 + 2c_i^T L^{-1}\nabla\hat{z}}{2\sigma_{i|I}}\right) - 1\right) \end{aligned} \tag{3.46}$$

The bias-affected success rate is a useful diagnostic tool in order to study the impact of biases on ambiguity resolution. For example one could study the impact of unmodeled ionospheric errors or multipath, see (Teunissen et al. 2000; Joosten et al. 2002).

3.3 The ambiguity residuals

The ambiguity residuals are defined as:

$$\check{\epsilon} = \hat{a} - \check{a} \tag{3.47}$$

These residuals can be used as a measure of the inconsistency between the data and the model. In order to measure the significance of the inconsistency also their probabilistic properties are needed. In practice it is often assumed that the fixed estimator $\check{a}$ is deterministic, so that the residuals simply take the same probability distribution as the float ambiguities, i.e. a normal distribution. This assumption, however, is not correct. It should only be used if the PDF of the float ambiguities is sufficiently peaked, so that the probability mass in S_a is very close to one.

3.3.1 Parameter distribution of the ambiguity residuals

In Teunissen (2002) and Verhagen and Teunissen (2004c) it was shown how the PDF of the ambiguity residuals can be constructed. Here a somewhat different approach will be followed.

Let $\mathrm{R} \subset \mathbb{R}^n$ be an arbitrary subset. Then

$$\begin{aligned}
P(\check{\epsilon} \in \mathrm{R}) &= P(\hat{a} - \check{a} \in \mathrm{R}) \\
&= \sum_{z \in \mathbb{Z}^n} P(\hat{a} - \check{a} \in \mathrm{R}, \hat{a} \in S_z) \\
&= \sum_{z \in \mathbb{Z}^n} P(\hat{a} - z \in \mathrm{R}, \hat{a} \in S_z) \\
&= \sum_{z \in \mathbb{Z}^n} P(\hat{a} \in \{\mathrm{R} + z\} \cap S_z) \\
&= \sum_{z \in \mathbb{Z}^n} \int_{\{\mathrm{R}+z\} \bigcap S_z} f_{\hat{a}}(x) dx \qquad \text{use change of variable } y = x - z \\
&= \sum_{z \in \mathbb{Z}^n} \int_{\mathrm{R} \bigcap S_0} f_{\hat{a}}(y + z) dy \\
&= \int_{\mathrm{R}} \sum_{z \in \mathbb{Z}^n} f_{\hat{a}}(x + z) s_0(x) dx
\end{aligned}$$

with $s_0(x)$ the indicator function, see equation (3.7). Since the above equation holds

for any R, it follows that the PDF of $\check{\epsilon}$ is given as:

$$\begin{aligned} f_{\check{\epsilon}}(x) &= \sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z)s_0(x) \\ &= \frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \sum_{z\in\mathbb{Z}^n} \exp\{-\frac{1}{2}\|x-a+z\|^2_{Q_{\hat{a}}}\}s_0(x) \end{aligned} \tag{3.48}$$

Note that $f_{\check{\epsilon}}(x)$ is only sensitive to the fractional part of a, that is the departure from integerness if $a \notin \mathbb{Z}^n$. But in general it is of course assumed that $a \in \mathbb{Z}^n$, and then a can be eliminated from equation (3.48) because of the infinite sum over all integers.

As was shown in Teunissen (2002) and Verhagen and Teunissen (2004c), the joint PDF of $\check{a}$ and $\check{\epsilon}$ is given by $f_{\check{\epsilon},\check{a}}(x,z) = f_{\hat{a}}(x+z)s_0(x)$. Figure 3.7 shows all steps required for the construction of the PDF of the ambiguity residuals in the one-dimensional (1-D) case. The PDF of $\hat{a}$ (top left) is plotted along the x-axis, the PMF of $\check{a}$ (top right) along the z-axis, and the joint PDF (top middle) is plotted in the xz-plane. Its construction from the marginal PDF and PMF can be seen as follows. First, the parts of the PDF of $\hat{a}$ are sliced out that correspond to all pull-in regions. For the 1-D case, the pull-in regions simply are intervals with length 1, centered at the integers, $S_z = \{x \in \mathbb{R} |\ |x-z| \leq \frac{1}{2}\}$. These slices are then translated along the z-axis to the corresponding integers z. The joint PDF of $\check{\epsilon}$ and $\check{a}$ (bottom left) follows from another translation of the slices, but now along the x-axis, so that they are all centered at the mean value $x = 0$. The PDF of $\check{\epsilon}$ (bottom right) is finally obtained by a summation over z, i.e. all slices are again translated along the z-axis to the origin.

Unfortunately, it is not possible to evaluate the PDF of the ambiguity residuals exactly because of the infinite sum over all integer vectors in $\mathbb{Z}^n$. It is therefore required to choose a finite subset $\Theta \subset \mathbb{Z}^n$ that will give a good approximation, see (Verhagen and Teunissen 2004b). This will be the case if:

$$\sum_{u\in\mathbb{Z}^n\setminus\Theta} \frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \exp\{-\frac{1}{2}\|x-a+u\|^2_{Q_{\hat{a}}}\} \approx 0 \tag{3.49}$$

In order to have a high probability that (3.49) holds, Θ should be chosen as:

$$\Theta = \{z \in \mathbb{Z}^n \mid \frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \exp\{-\frac{1}{2}\|x-a+z\|^2_{Q_{\hat{a}}}\} > \lambda\} \tag{3.50}$$

with λ very small, so that the contribution of all $u \in \mathbb{Z}^n\setminus\Theta$ to the infinite sum of equation (3.48) can be neglected. This holds true if the probability

$$P(\frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \exp\{-\frac{1}{2}\|x-a+z\|^2_{Q_{\hat{a}}}\} > \lambda)$$

is close to one. This probability can be computed as follows:

$$
\begin{aligned}
&P(\underbrace{\frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}}\exp\{-\frac{1}{2}\|x-a+z\|^2_{Q_{\hat{a}}}\}}_{=f_{\hat{a}}(x+z)} > \lambda) \\
&\qquad = P(\|x-a+z\|^2_{Q_{\hat{a}}} < \underbrace{-2\ln(\lambda\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n})}_{\chi^2}) \\
&\qquad = P(\|\hat{a}-a\|^2_{Q_{\hat{a}}} < \chi^2)
\end{aligned}
\tag{3.51}
$$

This probability can be computed since $\|\hat{a}-a\|^2_{Q_{\hat{a}}} \sim \chi^2(n,0)$, see appendix A.2.2.

Summarizing, this means that $f_{\check{\epsilon}}(x)$ can be approximated by:

$$
f_{\check{\epsilon}}(x) = \frac{1}{\sqrt{|Q_{\hat{a}}|}(2\pi)^{\frac{1}{2}n}} \sum_{z\in\Theta} \exp\{-\frac{1}{2}\|x+z\|^2_{Q_{\hat{a}}}\} s_0(x) \tag{3.52}
$$

with

$$
\Theta = \{z \in \mathbb{Z}^n \mid \|x+z\|^2_{Q_{\hat{a}}} < \chi^2\} \tag{3.53}
$$

where $a \in \mathbb{Z}^n$ is eliminated from equation (3.48), and χ^2 as defined in equation (3.51). So, the integer set contains all integers z within the ellipsoid centered at $-x$ and its size governed by χ.

3.3.2 PDF evaluation

Figure 3.9 shows $f_{\check{\epsilon}}(x)$ under H_0 for different values of σ in the one-dimensional case ($n = 1$). Also the extreme cases, $\sigma^2 = 0$ and $\sigma^2 \to \infty$, are shown. In the first case, an impulse PDF is obtained, in the second case a uniform PDF. Note that the unit of σ is $cycles$.

It can be seen that the PDF becomes peaked if the precision is better, i.e. $\sigma \downarrow 0$. In that case most of the probability mass of the PDF of $\hat{a}$ is located in the pull-in region S_a. This is the case for $\sigma = 0.1$. For $\sigma \geq 0.3$ (right panel) the distribution function becomes flat, and already for $\sigma = 1$ the PDF is very close to the uniform distribution.

Figure 3.10 shows the PDFs of $\hat{a}$, $\hat{z}$ and $\check{\epsilon}$ obtained for the three admissible estimators. The following vc-matrices (original and Z-transformed) were used:

$$
Q_{\hat{a}} = \begin{pmatrix} 4.9718 & 3.8733 \\ 3.8733 & 3.0188 \end{pmatrix}, \quad Q_{\hat{z}} = \begin{pmatrix} 0.0865 & -0.0364 \\ -0.0364 & 0.0847 \end{pmatrix} \tag{3.54}
$$

These vc-matrices are obtained for the dual-frequency geometry-free GPS model for one satellite-pair, with undifferenced code and phase standard deviations of 30 cm and 3 mm, respectively. The results were evaluated in Verhagen and Teunissen (2004c). Especially for the bootstrapped and the ILS estimator, the shape of the PDF of the residuals 'fits' the shape of the pull-in region quite well. The PDFs are multi-modal for

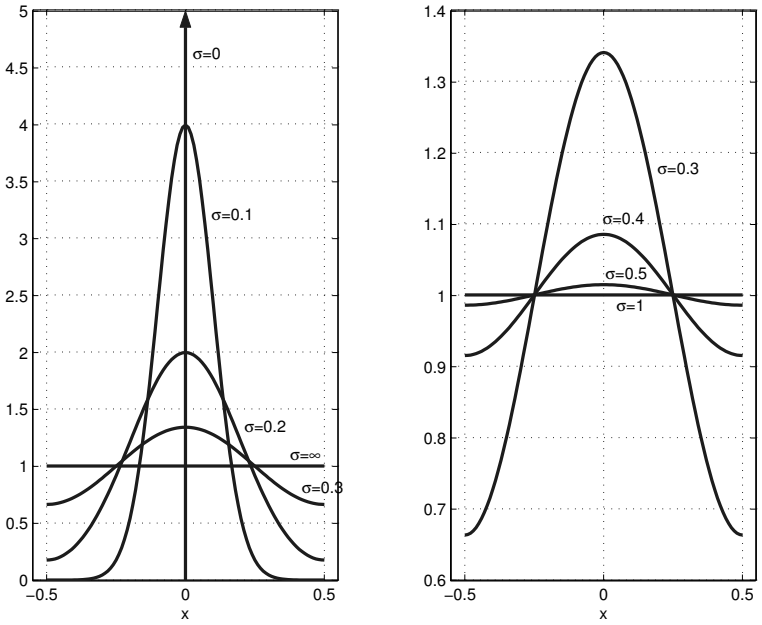

Figure 3.9: PDF of $\breve{\epsilon}$ for different values of σ.

the untransformed vc-matrices, but not for the Z-transformed vc-matrices. However, the shape of the PDF near the boundaries of the pull-in region is clearly different from the shape of the PDF of $\hat{a}$. This shows that in all cases the fixed ambiguity estimator should not be considered deterministic, since that would result in the assumption that $f_{\breve{\epsilon}}(x) := f_{\hat{a}}(x + \breve{a})$. This is only true if all the probability mass of $\hat{a}$ is located in the pull-in region S_a.

Figure 3.11 shows the ellipses that correspond to the vc-matrices of $\hat{a}$ and $\hat{z}$, and to the vc-matrices of the ambiguity residuals.

Approximation

In section 3.3.1 it was explained that only an approximation of the PDF of $\breve{\epsilon}$ is possible by replacing the infinite sum over all integers in equation (3.48) by a sum over a finite set of integers. It is investigated here how good the approximations work for different choices of λ, which determines the integer set, see equation (3.51). For that purpose, 10,000 samples of float ambiguities were generated using simulation for various vc-matrices, see appendix B. The corresponding ambiguity residuals were determined, so that the PDF of these parameters could be determined.

The results for the 2-D case are shown in figure 3.12. Besides $Q_{\hat{z}\_02\_01}$, also $\frac{1}{4}Q_{\hat{z}\_02\_01}$ and $4Q_{\hat{z}\_02\_01}$ were used in order to study the effect of higher/lower precision. The

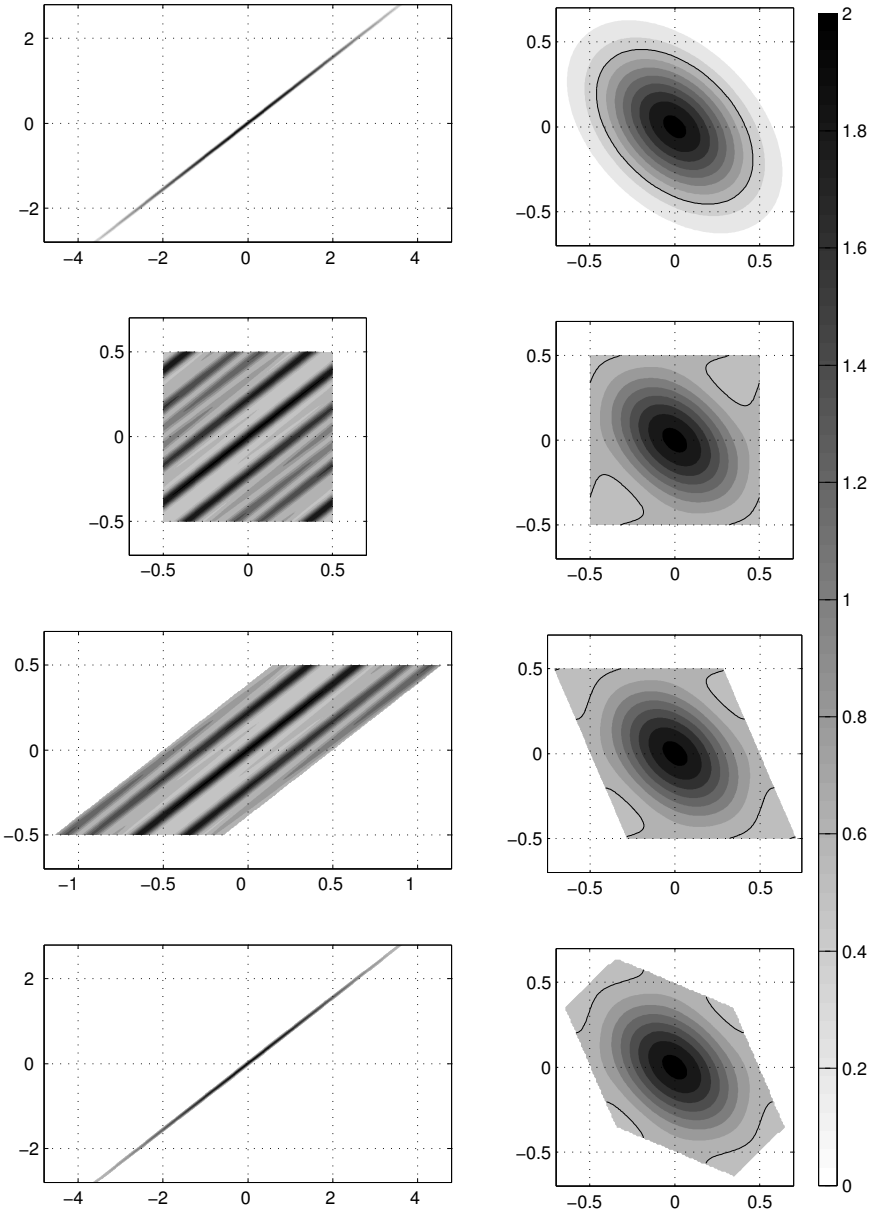

Figure 3.10: PDFs of $\hat{a}$, $\hat{z}$ and $\check{\epsilon}$. Left: for $Q_{\hat{a}}$; Right: for $Q_{\hat{z}}$. Top: PDF of $\hat{a}$ and $\hat{z}$; 2nd from top: PDF of $\epsilon_{\check{\mathsf{R}}}$; 3rd from top: PDF of $\epsilon_{\check{\mathsf{B}}}$; Bottom: PDF of $\epsilon_{\check{\mathsf{LS}}}$.

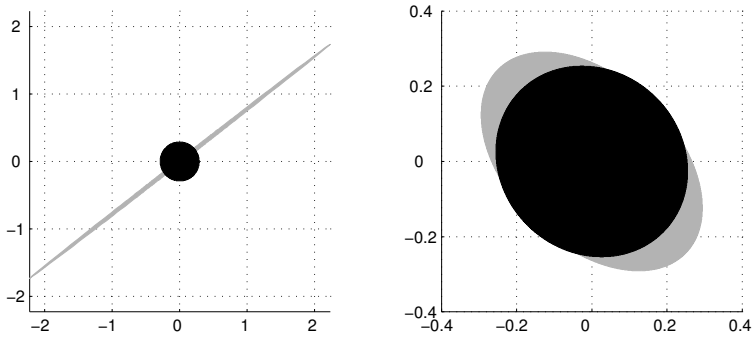

Figure 3.11: Ellipses corresponding to the vc-matrices of the float ambiguities (grey) and the ambiguity residuals (black). Left: for $Q_{\hat{a}}$; Right: for $Q_{\hat{z}}$.

horizontal axis gives the value of α:

$$\alpha = 1 - P(\|\hat{a} - a\|^2_{Q_{\hat{a}}} < \chi^2), \tag{3.55}$$

see equation (3.51).

The 'correct' values, $f^c_{\check{\epsilon}}(x)$, were computed for all samples using a very large integer set. Then the difference with the approximations $df = f^c_{\check{\epsilon}}(x) - f^\alpha_{\check{\epsilon}}(x)$ were computed. The top panels show along the vertical axes:

$$y = \frac{df}{N}, \qquad \text{(mean error)}$$

$$y = \arg\max_x(df), \qquad \text{(maximum error)}$$

where N is the number of samples.

It follows that for lower precision – top panels – a large integer set is required in order to get a good approximation. Note that the number of integers in the set Θ depends on the sample value of $\check{\epsilon}$, since the ellipsoidal region is centered at $-\check{\epsilon}$.

The results for the geometry-based examples are summarized in table 3.4. For each example, four different integer sets were chosen, based on the choice of α, see equation (3.55). The corresponding value of λ in equation (3.51) is also given. The largest of these sets was chosen such that the error in the approximation of the PDF was small enough to be neglected. It follows that the approximation errors have the same order as α. So, with $\alpha = 10^{-6}$ the errors are small, and the number of integers in the set Θ is also small enough to guarantee reasonable computation times.

3.4 Quality of the fixed baseline estimator

The distributional properties of the ambiguity parameters were discussed in the preceding sections. A user, however, will of course be mainly interested in the quality of the baseline solution. In this section, the distributional properties of the fixed baseline estimator will

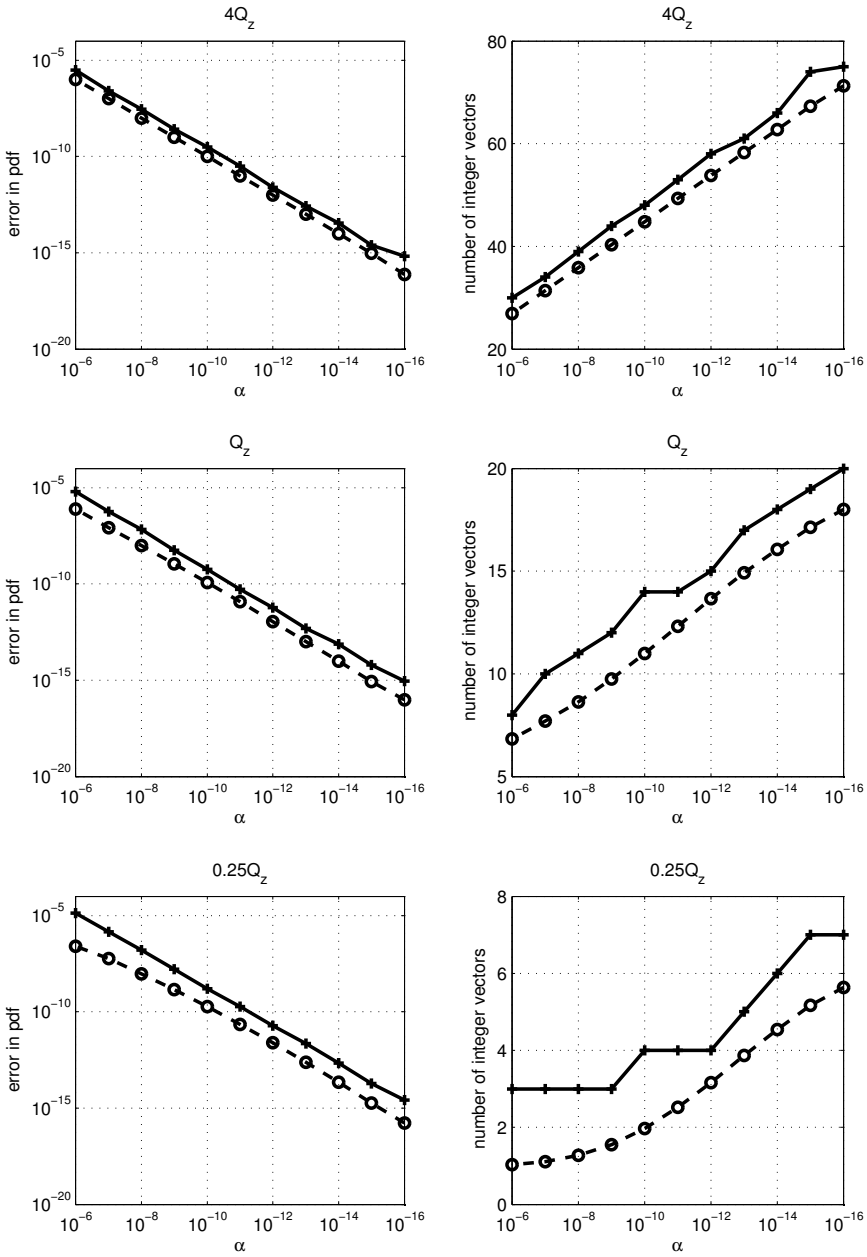

Figure 3.12: Approximation errors in $f_{\hat{\varepsilon}}(x)$. Top: $4Q_{\hat{z}\_02\_01}$; Center: $Q_{\hat{z}\_02\_01}$; Bottom: $\frac{1}{4}Q_{\hat{z}\_02\_01}$. Left: mean errors (dashed) and maximum errors (solid); Right: the mean number (dashed) and maximum number (solid) of integer vectors in integer set.

Table 3.4: Order of the mean and maximum approximation errors in $f_{\check{\epsilon}}(x)$. N_z equals the minimum and maximum number of integers used.

	α	λ	mean	maximum	N_z
06_01	10^{-16}	10^{-18}	10^{-16}	10^{-15}	753-850
	10^{-12}	10^{-14}	10^{-12}	10^{-12}	347-407
	10^{-8}	10^{-9}	10^{-8}	10^{-8}	112-155
	10^{-6}	10^{-7}	10^{-6}	10^{-6}	53-87
06_02	10^{-16}	10^{-19}	10^{-17}	10^{-15}	4730-4922
	10^{-12}	10^{-15}	10^{-12}	10^{-12}	2203-2379
	10^{-8}	10^{-10}	10^{-8}	10^{-8}	785-879
	10^{-6}	10^{-8}	10^{-6}	10^{-6}	357-442
10_01	10^{-16}	10^{-18}	10^{-16}	10^{-14}	274-345
	10^{-12}	10^{-14}	10^{-12}	10^{-12}	80-121
	10^{-8}	10^{-9}	10^{-8}	10^{-8}	11-31
	10^{-6}	10^{-6}	10^{-6}	10^{-6}	2-14
10_03	10^{-16}	10^{-20}	10^{-17}	10^{-15}	44323-45452
	10^{-12}	10^{-16}	10^{-12}	10^{-12}	14535-15048
	10^{-8}	10^{-11}	10^{-8}	10^{-8}	2959-3330
	10^{-6}	10^{-9}	10^{-6}	10^{-6}	1007-1215

be presented. Furthermore, some baseline probabilities will be given that can be used to obtain a quality assessment of the fixed baseline solution.

3.4.1 Parameter distributions of the baseline estimators

The float baseline solution is assumed to be normally distributed, i.e. $\hat{b} \sim N(b, Q_{\hat{b}})$. The PDF of $\hat{b}$ reads then

$$f_{\hat{b}}(x) = \frac{1}{\sqrt{|Q_{\hat{b}}|}(2\pi)^{\frac{1}{2}p}} \exp\{-\frac{1}{2}\|x-b\|^2_{Q_{\hat{b}}}\} \tag{3.56}$$

The conditional PDF is given by:

$$f_{\hat{b}|\hat{a}}(x|z) = \frac{1}{\sqrt{|Q_{\hat{b}|\hat{a}}|}(2\pi)^{\frac{1}{2}p}} \exp\{-\frac{1}{2}\|x-b(z)\|^2_{Q_{\hat{b}|\hat{a}}}\} \tag{3.57}$$

with conditional mean $b(z) = b - Q_{\hat{b}\hat{a}}Q^{-1}_{\hat{a}}(a-z)$, and conditional variance matrix $Q_{\hat{b}|\hat{a}} = Q_{\hat{b}} - Q_{\hat{b}\hat{a}}Q^{-1}_{\hat{a}}Q_{\hat{a}\hat{b}}$.

The distribution function of the fixed baseline parameters, $f_{\check{b}}(x)$, was derived in Teunis-

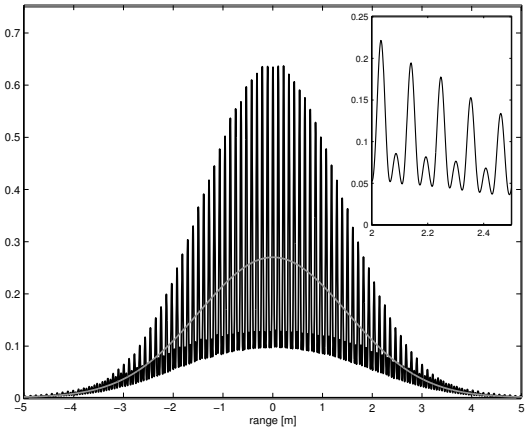

Figure 3.13: Example of the multi-modal PDF of $\check{b}$ and the corresponding uni-modal PDF of $\hat{b}$.

sen (1998e;2002) as:

$$f_{\check{b}}(x) = \sum_{z \in \mathbb{Z}^n} f_{\hat{b}|\hat{a}}(x|z) P(\check{a} = z) \tag{3.58}$$

So, although the model (3.1) is linear and the observations are normally distributed, the distribution of the fixed baseline solution is not normal, but multi-modal. Since $\hat{b}(z) = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(\hat{a} - z)$ it follows that the peaks are centered at $x = Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} z$, $z \in \mathbb{Z}^n$ and their height depends on $P(\check{a} = z)$. Figure 3.13 shows an example of the multi-modal PDF of $\check{b}$ and the corresponding PDF of $\hat{b}$.

From equations (3.4) and (3.30) it follows that the fixed baseline estimator is unbiased:

$$E\{\check{b}\} = b - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(a - E\{\check{a}\}) = b \tag{3.59}$$

When the propagation law of covariances is applied to equation (3.4), it follows that

$$Q_{\check{b}} = Q_{\hat{b}} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} Q_{\hat{a}\hat{b}} + Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} Q_{\check{a}} Q_{\hat{a}}^{-1} Q_{\hat{a}\hat{b}} \tag{3.60}$$

$$= Q_{\hat{b}|\hat{a}} + Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} Q_{\check{a}} Q_{\hat{a}}^{-1} Q_{\hat{a}\hat{b}} \tag{3.61}$$

So, only if $Q_{\check{a}} = 0$ the precision of the fixed baseline estimator is equal to that of the conditional baseline estimator. Equation (3.32) shows that $Q_{\check{a}} \approx 0$ if the probability that a wrong integer is estimated can be neglected.

If the ambiguities are fixed incorrectly, this may result in a fixed baseline solution that is worse than the float solution. In section 3.2.2 it was explained that it is therefore important to check whether or not one can have enough confidence in the correctness of the fixed ambiguity solution. But this does still not tell anything on the quality of the fixed baseline.

3.4.2 Baseline probabilities

The quality of the fixed baseline solution could be evaluated by means of the probability that $\check{b}$ lies in a certain convex region $E_b \subset \mathbb{R}^p$ symmetric with respect to b. This probability follows from equation (3.58) as:

$$P(\check{b} \in E_b) = \sum_{z \in \mathbb{Z}^n} \int_{E_b} f_{\hat{b}|\hat{a}}(x|z) dx P(\check{a} = z) \tag{3.62}$$

The infinite sum in the equation cannot be evaluated exactly and therefore in practice a lower and upper bound can be used, (Teunissen 1999b):

$$P(\hat{b}(a) \in E_b) P(\check{a} = a) \leq P(\check{b} \in E_b) \leq P(\hat{b}(a) \in E_b) \tag{3.63}$$

These bounds become tight when the success rate $P(\check{a} = a)$ is close to one. In that case also the following is true:

$$P(\check{b} \in E_b) \approx P(\hat{b}(a) \in E_b) \geq P(\hat{b} \in E_b) \tag{3.64}$$

The vc-matrix of the conditional baseline estimator, $Q_{\hat{b}|\hat{a}}$ is often used as a measure of the fixed baseline precision. Therefore, this vc-matrix will be used in order to define the shape of the confidence region E_b, so that it takes the ellipsoidal form:

$$E_b = \left\{ x \in \mathbb{R}^n \mid (x-b)^T Q_{\hat{b}|\hat{a}}^{-1} (x-b) \leq \beta^2 \right\} \tag{3.65}$$

The size of the region can thus be varied by the choice of β. Then:

$$P(\check{b} \in E_b) = P(\|\check{b} - b\|^2_{Q_{\hat{b}|\hat{a}}} \leq \beta^2) = \sum_{z \in \mathbb{Z}^n} P(\chi^2(p, \lambda_z) \leq \beta^2) P(\check{a} = z) \tag{3.66}$$

with $\chi^2(p, \lambda_z)$ the non-central χ^2-distribution with p degrees of freedom and non-centrality parameter λ_z:

$$\lambda_z = \|\nabla \check{b}_z\|^2_{Q_{\hat{b}|\hat{a}}}, \quad \nabla \check{b}_z = Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (z - a) \tag{3.67}$$

The interval of (3.63) becomes now:

$$\alpha_1 \leq P(\|\check{b} - b\|^2_{Q_{\hat{b}|\hat{a}}} \leq \beta^2) \leq \alpha_2 \tag{3.68}$$

with

$$\alpha_1 = \alpha_2 P(\check{a} = a) \quad \text{and} \quad \alpha_2 = P(\chi^2(p, 0) \leq \beta^2)$$

By choosing the confidence level α_1, the success rate can be used to determine α_2, which finally can be used to compute β with the aid of the χ^2-distribution.

Equation (3.68) is useful in order to decide whether or not one can have enough confidence in the fixed solution, although it will be difficult to know what value of β is

acceptable. Moreover, the bounds will only be tight if the success rate is high. Furthermore, it is also interesting to compare the 'quality' of the float and the fixed baseline estimators. For that purpose, the probabilities $P(\check{b} \in E_b)$ and $P(\hat{b} \in E_b)$ can be considered, with E_b identical in both cases. But then the choice of the region E_b is not so straightforward; with equation (3.65) it is not possible to give an exact evaluation of $P(\hat{b} \in E_b)$. Another option is therefore to consider the probabilities:

$$P(\|\hat{b} - b\|^2_{Q_{\hat{b}}} \leq \beta^2) \quad \text{and} \quad P(\|\check{b} - b\|^2_{Q_{\check{b}}} \leq \beta^2) \tag{3.69}$$

The first probability can be evaluated exactly since $\|\hat{b}-b\|^2_{Q_{\hat{b}}}$ has a central χ^2-distribution. The second probability, however, cannot be evaluated exactly.

Until now the quality of the complete baseline estimator was considered. However, a user will be mainly interested in the baseline increments, whereas the baseline estimator may contain other parameters as well, such as ionospheric delays. Therefore, it is more interesting to look at the probabilities

$$P(\|\hat{b}_x - b_x\| \leq \beta) \quad \text{and} \quad P(\|\check{b}_x - b_x\| \leq \beta) \tag{3.70}$$

where the subscript x is used to indicate that only the baseline parameters referring to the actual receiver position are considered, i.e. $b_x = \mathbf{r}_{qr}$ from equation (2.39). So, $\|\hat{b}_x - b_x\|$ is the actual distance to the true receiver position.

Examples

It is possible to evaluate the baseline probabilities using simulations. Examples 06_01 and 06_02 from appendix B will be considered here. A large number of float baseline estimates was generated, and for each float sample the corresponding fixed baseline was computed. The vc-matrix of the fixed baseline estimator was determined from the samples. Figure 3.14 shows the various baseline probabilities. It can be seen that the bounds of equation (3.68) are generally not strict (shown in grey), and it is difficult to interpret the probabilities with E_b chosen as in equation (3.65). More examples and an evaluation of the results will be given in chapters 4 and 5, when the performance of the float and fixed estimators is compared to that of new estimators introduced in those chapters.

3.5 Validation of the fixed solution

A parameter estimation theory cannot be considered complete without rigorous measures for validating the parameter solution. In the classical theory of linear estimation, the vc-matrices provide sufficient information on the precision of the estimated parameters. The reason is that a linear model applied to normally distributed (Gaussian) data, provides linear estimators that are also normally distributed, and the peakedness of the multivariate normal distribution is completely captured by the vc-matrix.

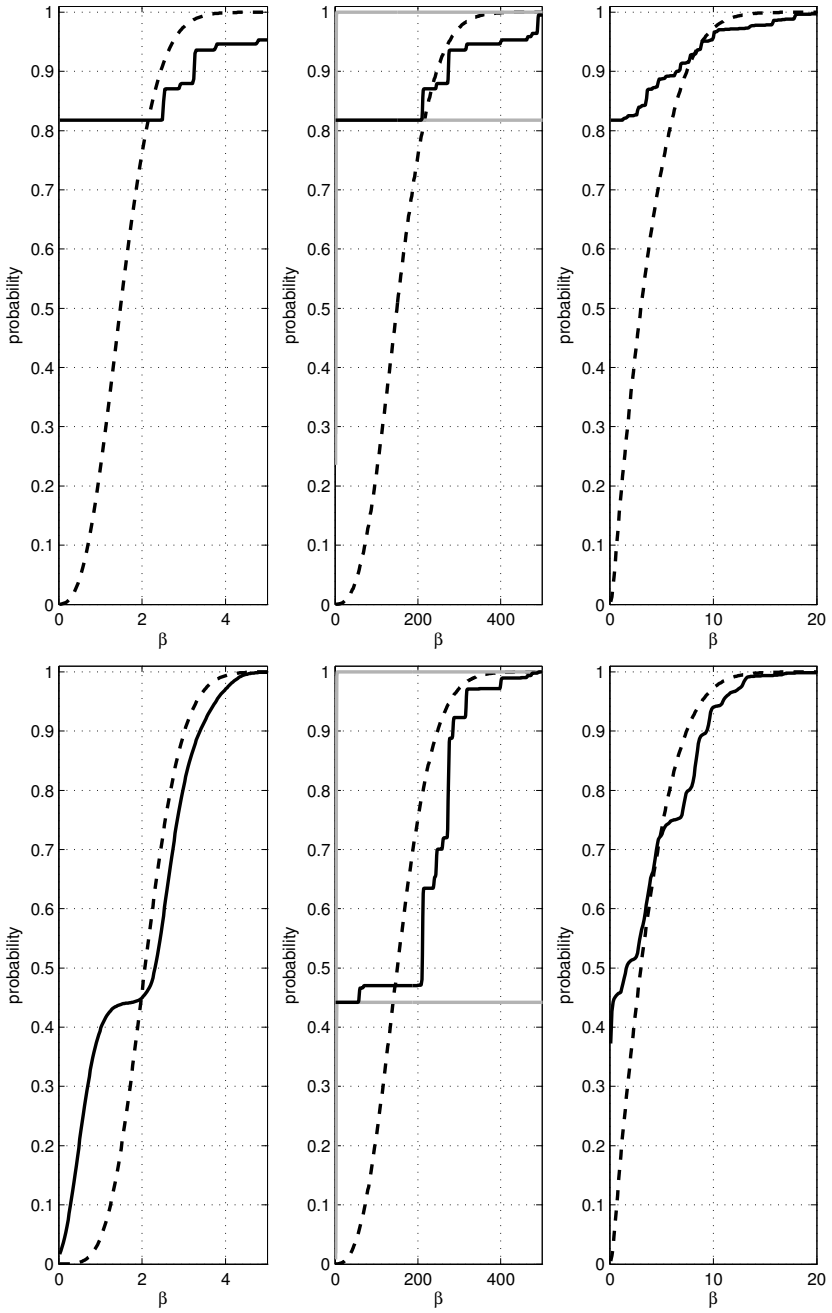

Figure 3.14: Float (dashed) and fixed (solid) baseline probabilities for examples 06_01 (top) and 06_02 (bottom).
Left: $P(\|\hat{b}-b\|^2_{Q_{\hat{b}}} \leq \beta^2)$ and $P(\|\check{b}-b\|^2_{Q_{\check{b}}} \leq \beta^2)$.
Center: $P(\|\hat{b}-b\|^2_{Q_{\hat{b}|\hat{a}}} \leq \beta^2)$ and $P(\|\check{b}-b\|^2_{Q_{\check{b}|\hat{a}}} \leq \beta^2)$.
Right: $P(\|\hat{b}_x-b_x\| \leq \beta)$ and $P(\|\check{b}_x-b_x\| \leq \beta)$.

Unfortunately, this relatively simple approach cannot be applied when integer parameters are involved in the estimation process, since the integer estimators do not have a Gaussian distribution, even if the model is linear and the data are normally distributed. Instead of the vc-matrices, the parameter distribution itself has to be used in order to obtain the appropriate measures that can be used to validate the integer parameter solution.

In the past, the problem of non-Gaussian parameter distributions was circumvented by simply ignoring the randomness of the fixed ambiguity estimator. Several testing procedures for the validation of the fixed solution have been proposed using this approach. This section will give an overview and evaluation of these procedures, taken from (Verhagen 2004).

3.5.1 Integer validation

The first step in validating the float and fixed solution is to define three classes of hypotheses, (Teunissen 1998b). In practice the following hypotheses are used:

$$\begin{aligned} H_1 &: y = Aa + Bb + e, && a \in \mathbb{R}^n,\ b \in \mathbb{R}^p,\ e \in \mathbb{R}^m \\ H_2 &: y \in \mathbb{R}^m, && \\ H_3 &: y = A\check{a} + Bb + e, && b \in \mathbb{R}^p,\ e \in \mathbb{R}^m \end{aligned} \tag{3.71}$$

with $Q_y = \sigma^2 G_y$ for all hypotheses.

In hypothesis H_1 it is assumed that all unknown parameters are real-valued. In other words, the constraint that the ambiguities should be integer-valued is ignored. The float solution is thus the result of least-squares estimation under H_1.

In the third hypothesis the integer constraint is considered by assuming that the correct integer values of the ambiguities are known. The values are chosen equal to the fixed ambiguities obtained with integer least-squares. Hence, the least-squares solution under H_3 corresponds to the fixed solution. Clearly, the first hypothesis is more relaxed than the third hypothesis, i.e. $H_3 \subset H_1$.

The second hypothesis does not put any restrictions on y and is thus the most relaxed hypothesis.

It will be assumed that the m-vector of observations y is normally distributed with a zero-mean residual vector e. The matrix G_y is the cofactor matrix of the variance-covariance matrix Q_y, and σ^2 is the variance factor of unit weight. The unbiased estimates of σ^2 under H_1 and H_3 respectively are given by:

$$\begin{aligned} H_1 &: \hat{\sigma}^2 = \frac{\hat{e}^T G_y^{-1} \hat{e}}{m-n-p} = \frac{\hat{\Omega}}{m-n-p} \\ H_3 &: \check{\sigma}^2 = \frac{\check{e}^T G_y^{-1} \check{e}}{m-p} = \frac{\check{\Omega}}{m-p} \end{aligned} \tag{3.72}$$

The denominators in (3.72) are equal to the redundancies under the corresponding hypotheses.

Since $H_3 \subset H_1 \subset H_2$, starting point is to test H_1 against H_2. This will answer the question whether or not the model on the basis of which the float solution is computed may be considered valid or not. This is important, since the data may be contaminated by undetected errors and/or the model may not represent reality because some physical or geometrical effects are not captured correctly (e.g. multipath, atmospheric delays).

The test statistic that allows to test H_1 against H_2 is given by $\frac{\hat{\sigma}^2}{\sigma^2}$, and H_1 is accepted when

$$\frac{\hat{\sigma}^2}{\sigma^2} < F_\alpha(m-n-p, \infty, 0) \tag{3.73}$$

where $F_\alpha(m-n-p, \infty, 0)$ is the critical value that corresponds to the central F-distribution, $F(m-n-p, \infty, 0)$, and the chosen level of significance α. This test is identical to the overall model test in section 2.5.2.

If the value of the test statistic has passed the test in expression (3.73), the next step is to verify whether or not hypothesis H_3 may be considered valid. This means testing H_3 against H_1. Alternatively, one could test H_3 against H_2 in the case one is not sure if the first hypothesis is true. This would result in the following test:

$$\frac{\check{\sigma}^2}{\sigma^2} < F_\alpha(m-p, \infty, 0) \tag{3.74}$$

However, testing H_3 against H_1 is more powerful and focuses on the question whether the fixed ambiguity vector $\check{a}$ may be considered valid. The test statistic is given by $\frac{(\hat{a}-\check{a})^T G_{\hat{a}}^{-1}(\hat{a}-\check{a})}{n\sigma^2}$, which is distributed under H_3 as $F(n, \infty, 0)$. Hence, the fixed ambiguities are considered valid if:

$$\frac{(\hat{a}-\check{a})^T G_{\hat{a}}^{-1}(\hat{a}-\check{a})}{n\sigma^2} = \frac{R_1}{n\sigma^2} < F_\alpha(n, \infty, 0) \tag{3.75}$$

Note that from equation (3.14), with the last term equal to zero, it follows that:

$$R_1 = (\hat{a}-\check{a})^T G_{\hat{a}}^{-1}(\hat{a}-\check{a}) = \check{e}^T G_y^{-1}\check{e} - \hat{e}^T G_y^{-1}\hat{e} = \check{\Omega} - \hat{\Omega}$$

This shows that test (3.75) can also be expressed in terms of the test statistics of (3.73) and (3.74).

The test in (3.75) is based on the distance - as measured by the metric $Q_{\hat{a}}$ - between the float and fixed solution. This is reasonable, since one would not have much confidence in the fixed solution if the distance to the float solution is large. In that case one should not use the fixed solution and instead stick to the float solution and/or gather more data so that the estimation and validation process can be repeated.

The value of the variance factor is not always known. In that case σ^2 can be replaced by $\hat{\sigma}^2$, so that the test becomes:

$$\frac{R_1}{n\hat{\sigma}^2} < F_\alpha(n, m-n-p, 0) \tag{3.76}$$

The means of the test statistics of (3.73), (3.74) and (3.75) are all equal to one, whereas the mean of the test statistic of (3.76) equals $(m-n-p)/(m-n-p-2)$, which is

always larger than one. This shows that the distributions of the tests (3.76) and (3.75) approach each other in the case $m-n-p$ is large. Note that the test statistic in equation (3.76) is assumed to have an F-distribution, which is only the case if $\hat{\Omega}$ and R_1 are independent. In general this is not true.

3.5.2 Discrimination tests

The tests of (3.75) and (3.76) do not give information on the likelihood of the fixed solution as compared to another set of integers. There might be another integer vector, say $\breve{a}'$, that results in a value for the test statistic that passes test (3.75). Of course it is known that the likelihood of the fixed solution $\breve{a}$ is always larger than the likelihood of any other integer vector in the case it was obtained with integer least-squares estimation. However, if the likelihood of $\breve{a}$ is not sufficiently larger than the likelihood of $\breve{a}'$, the two solutions cannot be discriminated with enough confidence. Therefore, the ratio of $\breve{\sigma}'^2/\breve{\sigma}^2$ seems an intuitively appealing test statistic. If this ratio is significantly larger than one, $\breve{a}$ is considered far more likely than $\breve{a}'$. So, $\breve{a}$ is validated against any other $\breve{a}'$ when:

$$\frac{\breve{\sigma}'^2}{\breve{\sigma}^2} = \frac{\breve{\Omega}'}{\breve{\Omega}} \geq k \tag{3.77}$$

in which $k > 1$ is the chosen critical value. This test is referred to as the F-ratio test. It only needs to be carried out with $\breve{a}'$ chosen as the second-best integer estimate obtained with integer least-squares, which will be denoted as $\breve{a}_2$ with corresponding $\breve{\sigma}_2^2$. If the test is passed in that case, it will pass for any other choice of $\breve{a}'$.

Note that the choice of the critical value is not trivial. In practice, it is often assumed that the test statistic has an F-distribution, so that the critical value can be based on a choice of the level of significance. However, as mentioned in Teunissen (1998b) this is not correct since $\breve{\sigma}_2^2$ and $\breve{\sigma}^2$ are not independent. Still, the test in (3.77) is often used and seems to work satisfactorily, see e.g. (Frei and Beutler 1990; Abidin 1993). In Landau and Euler (1992) it is proposed to use the same test statistic but with a critical value of 2.

Variations of test (3.77) are also proposed in literature. For example, Euler and Schaffrin (1991) use the so-called ratio test given by:

$$\frac{R_2}{R_1} \geq k \tag{3.78}$$

with $R_2 = (\hat{a} - \breve{a}_2)^T G_{\hat{a}}^{-1} (\hat{a} - \breve{a}_2)$. This ratio test is derived by applying the classical theory of hypothesis testing. Three hypotheses are considered:

$$H_0 : \; a = \hat{a}, \quad H_1 : \; a = \breve{a} \quad H_2 : \; a = \breve{a}_2 \tag{3.79}$$

The following test statistics are then defined:

$$T_i = \frac{R_i}{n\hat{\sigma}^2} \sim F(n, m-n-p, \lambda_i), \; i = 1, 2 \tag{3.80}$$

where $F(n, m-n-p, \lambda_i)$ denotes the non-central F-distribution with non-centrality parameter λ_i.

The probability of incorrectly accepting H_0 when in fact H_i is true, β_i, is given by:

$$\beta_i = \int_0^{F_\alpha(n,m-n-p,0)} f_{T_i}(x \mid n, m-n-p, \lambda_i)dx \tag{3.81}$$

where $f_{T_i}(x|n, m-n-p, \lambda_i)$ is the PDF of T_i and $F_\alpha(n, m-n-p, 0)$ is the critical value of a given false alarm rate, α.

By choosing α, β_1, and β_2, the corresponding non-centrality parameters, $\lambda_i(\alpha, \beta_i)$, can be determined. Finally, the critical value of test (3.78) is then determined as:

$$k = \frac{\lambda_2(\alpha, \beta_2)}{\lambda_1(\alpha, \beta_1)} \tag{3.82}$$

However, the determination of the non-centrality parameters such that the probabilities β_i are obtained, is not straightforward. Moreover, the choice of these probabilities is not obvious. Euler and Schaffrin (1991) used test computations, from which followed that a critical value between 5 and 10 should be chosen depending on the degrees of freedom.

In literature it has also been proposed to use a fixed critical value instead of using equation (3.82).

Wei and Schwarz (1995) propose to use the test with a critical value of $k = 2$. They do not claim that there is a theoretical justification for using this test. Moreover, they acknowledge that it is risky to apply the test to an individual epoch in the case of poor satellite geometry or noisy observations. Therefore, it is proposed to apply the test to all epochs in a time period of 30-60 seconds, and only accept the integer solution if the same solution is accepted for all epochs. It is mentioned that the test is conservative.

Han and Rizos (1996a) showed that good results can be obtained with the ratio test with a critical value of $k = 1.5$ if one can have confidence in the correctness of the stochastic model. For their experiments this was accomplished by applying satellite elevation-dependent weighting, with the weights determined for the receiver at hand.

In many software packages a fixed value for the ratio is used, e.g. $k = 3$ (Leick 2003).

An alternative test, proposed by e.g. (Han and Rizos 1996b), is to combine test (3.76) with the following test:

$$\frac{R_2}{n\hat{\sigma}^2} > F_{\alpha'}(n, m-n-p, 0) \geq F_\alpha(n, m-n-p, 0) \tag{3.83}$$

Only if both tests are passed $\breve{a}$ is accepted. For that purpose $F_{\alpha'}(n, m-n-p, 0)$ should be chosen sufficiently larger than $F_\alpha(n, m-n-p, 0)$ in order to guarantee that $\breve{a}_2$ is much less likely than $\breve{a}$.

In Chen (1997) a similar approach is followed, but here the test is defined as:

$$\breve{\Omega}_2 > \chi^2_\alpha(m-p, 0) \quad \text{and} \quad \breve{\Omega} < \chi^2_\beta(m-p, \delta) \tag{3.84}$$

where $\chi^2_\alpha(m-p,0)$ is the critical value corresponding to the central χ^2-distribution with level of significance α (= probability in right-hand tail); $\chi^2_\beta(m-p,\delta)$ is the critical value corresponding to the non-central χ^2-distribution with probability β (= probability in left-hand tail) and non-centrality parameter:

$$\delta = (\check{a} - \check{a}_2)^T G_{\hat{a}}^{-1} (\check{a} - \check{a}_2) \tag{3.85}$$

Still another approach would be to look at the difference of the quadratic forms R_1 and R_2, see (Tiberius and De Jonge 1995). First, test (3.76) is carried out. When it is passed, the next step is to perform the difference test:

$$R_2 - R_1 \geq \sigma^2 k \tag{3.86}$$

It will be clear, that again the problem with this test is the choice of the critical value k, only an empirically determined value can be used.

Another test worth mentioning is the one proposed in Wang et al. (1998):

$$\frac{\check{\Omega}_2 - \check{\Omega}}{\sqrt{\hat{\sigma}^2 2\sqrt{\delta}}} > t_\alpha(m-n-p) \tag{3.87}$$

where $t_\alpha(m-n-p)$ is the critical value corresponding to the Student's t-distribution, see appendix A.2.4. In order to arrive at this result it is stated that the difference $d = \check{\Omega}_2 - \check{\Omega} = R_2 - R_1$ has a truncated normal distribution. This is based on the following derivation:

$$\begin{aligned} d &= \check{\Omega}_2 - \check{\Omega} \\ &= 2(\check{a} - \check{a}_2)^T Q_{\hat{a}}^{-1} \hat{a} + \check{a}_2^T Q_{\hat{a}}^{-1} \check{a}_2 - \check{a}^T Q_{\hat{a}}^{-1} \check{a} \\ &= 2(\check{a} - \check{a}_2)^T Q_{\hat{a}}^{-1} \begin{pmatrix} Q_{\hat{b}\hat{a}} & Q_{\hat{a}} \end{pmatrix} A^T Q_y^{-1} y + \check{a}_2^T Q_{\hat{a}}^{-1} \check{a}_2 - \check{a}^T Q_{\hat{a}}^{-1} \check{a} \\ &= d_1 A^T Q_y^{-1} y + d_0 \end{aligned}$$

It is then assumed that the terms d_1 and d_0 are deterministic. This, however, is not true, since both $\check{a}$ and $\check{a}_2$ depend on y. Therefore, the test statistic of (3.87) does actually not have the t-distribution.

In Han (1997) yet another test was proposed. The null hypothesis that is used is H_3 in equation (3.71), the alternative hypothesis is defined as:

$$H_a : y = A\check{a} + A(\check{a}_2 - \check{a})\gamma + Bb + e,\ \gamma \in \mathbb{R},\ b \in \mathbb{R}^p,\ e \in \mathbb{R}^m$$

Testing the alternative hypothesis against the null hypothesis in this case means testing how likely it is that an outlier in the direction of $(\check{a}_2 - \check{a})$ occurred. The definition of the test statistic is based on the assumption that:

$$\frac{\gamma}{\hat{\sigma}_\gamma} \sim t(m-p-1) \tag{3.88}$$

with:

$$\hat{\gamma} = \frac{(\check{a}_2 - \check{a})^T G_{\hat{a}}^{-1} (\hat{a} - \check{a})}{\delta} \quad \text{and} \quad \hat{\sigma}_\gamma^2 = \frac{\check{\Omega} - \hat{\gamma}^2 \delta}{m-p-1} \tag{3.89}$$

The test is then defined by the constraint that $P(\gamma < 0.5 \mid H_a) > \beta$. However, the problem is again that the test statistic does actually not have a Student's t-distribution for the same reason as described above for test (3.87). Moreover, it can be easily shown that $\hat{\gamma}$ is always smaller than or equal to 0.5 so that the definition of the test is not rigorous:

$$\begin{aligned}\hat{\gamma} &= \frac{(\breve{a}_2 - \breve{a})^T G_{\hat{a}}^{-1}(\hat{a} - \breve{a})}{(\breve{a} - \breve{a}_2)^T G_{\hat{a}}^{-1}(\breve{a} - \breve{a}_2)} \\ &= \frac{\breve{a}_2^T G_{\hat{a}}^{-1}\hat{a} - \breve{a}^T G_{\hat{a}}^{-1}\hat{a} - \breve{a}_2^T G_{\hat{a}}^{-1}\breve{a} - \breve{a}^T G_{\hat{a}}^{-1}\breve{a}}{\breve{a}_2^T G_{\hat{a}}^{-1}\breve{a}_2 - 2\breve{a}_2^T G_{\hat{a}}^{-1}\breve{a} + \breve{a}^T G_{\hat{a}}^{-1}\breve{a}} \\ &= \frac{R_1 + \Gamma}{R_1 + R_2 + 2\Gamma} \\ &\leq \frac{R_1 + \Gamma}{2(R_1 + \Gamma)} = \frac{1}{2}\end{aligned} \tag{3.90}$$

with $\Gamma = \breve{a}_2^T G_{\hat{a}}^{-1}\hat{a} + \breve{a}^T G_{\hat{a}}^{-1}\hat{a} - \breve{a}_2^T G_{\hat{a}}^{-1}\breve{a} - \hat{a}^T G_{\hat{a}}^{-1}\hat{a}$.

It should be noted that in Wang et al. (1998) another test statistic was presented with a similar form as that of Han (1997).

3.5.3 Evaluation of the test statistics

A problem with all of the above-mentioned tests is the choice of the critical values. Either an empirically determined value has to be used, or they are based on the incorrect assumption that the fixed ambiguity estimator $\breve{a}$ may be considered deterministic. In principle, this is not true. Firstly, because the entries of the vector $\breve{a}$ depend on the same vector y as used in the formulation of the hypothesis H_3 in (3.71). So, if y changes, also $\breve{a}$ will change. Secondly, since the vector y is assumed to be random, also the fixed ambiguities obtained with integer least-squares estimation will be stochastic.

Integer validation based on for example the ratio tests in (3.77) and (3.78) often work satisfactorily. The reason is that the stochasticity of $\breve{a}$ may indeed be neglected if there is sufficient probability mass located at one integer grid point of $\mathbb{Z}^n$, i.e if the success rate is very close to one. So, one could use the additional constraint that the success rate should exceed a certain limit. But this immediately shows that this will not be the case in many practical situations. In those cases, the complete distribution function of the fixed ambiguities should be used for proper validation. It would thus be better to define an integer test that takes into account the randomness of the fixed ambiguities. Starting point is to replace the third hypothesis in (3.71) with the following, see (Teunissen 1998b):

$$H_3 : y = Aa + Bb + e, \quad a \in \mathbb{Z}^n, b \in \mathbb{R}^p, e \in \mathbb{R}^m \tag{3.91}$$

This hypothesis is more relaxed than the one used in (3.71), but makes it possible to find a theoretically sound criterion to validate the integer solution.

Another point of criticism is that the combined integer estimation and validation solution lacks an overall probabilistic evaluation.

Table 3.5: Overview of all test statistics to validate the integer ambiguity solution as proposed in literature. The fixed solution is accepted if both the integer test and the discrimination test are passed.

integer test	discrimination test	equation
$\frac{R_1}{n\hat{\sigma}^2} < F_\alpha(n, m-n-p)$	$\frac{\check{\Omega}_2}{\check{\Omega}} > F_{\alpha'}(m-p, m-p)$	(3.77)
$\frac{R_1}{n\hat{\sigma}^2} < F_\alpha(n, m-n-p)$	$\frac{R_2}{R_1} > k$	(3.78)
$\frac{R_1}{n\hat{\sigma}^2} < F_\alpha(n, m-n-p)$	$\frac{R_2}{n\hat{\sigma}^2} > F_{\alpha'}(n, m-n-p)$	(3.83)
$\check{\Omega} < \chi^2_\beta(m-p, \delta)$	$\check{\Omega}_2 > \chi^2_\alpha(m-p)$	(3.84)
$\frac{R_1}{n\hat{\sigma}^2} < F_\alpha(n, m-n-p)$	$R_2 - R_1 > k$	(3.86)
$\frac{R_1}{n\hat{\sigma}^2} < F_\alpha(n, m-n-p)$	$\frac{R_2-R_1}{2\sqrt{\hat{\sigma}^2\delta}} > t_{\alpha'}(m-n-p)$	(3.87)

An overview of all test statistics described in this section is given in table 3.5. It can be seen that there are three pairs of tests which have a similar 'structure': the two ratio tests of (3.77) and (3.78); two tests (3.83) and (3.84) that consider the best and second-best integer solutions separately; and the two tests (3.86) and (3.87) that look at the difference $d = R_2 - R_1$. Note, however, that the discrimination tests of (3.78) and (3.86) have an important conceptual difference as compared to the other tests. These two tests take only the float and fixed *ambiguity* solutions into account for discrimination, whereas the other tests take the complete float solution into account together with the fixed ambiguities.

It would be interesting to know which of the approaches works best in all possible situations. For this purpose in Verhagen (2004) simulations were used for a two-dimensional example (geometry-free model with two unknown ambiguities). It followed that the ratio test (3.78) and the difference test (3.86) perform well, provided that an appropriate critical value is chosen. However, that is exactly one of the major problems in practice.

It should be noted that in practice the ratio tests are often applied without the integer test (3.76), as it is thought that if one can have enough confidence in the float solution, the discrimination test is sufficient to judge whether or not the likelihood of the corresponding fixed ambiguity solution is large enough.

In chapter 5 it will be shown that it is possible to give a theoretical foundation for some of the discrimination tests. The performance of the tests will also be evaluated, and it will be shown that one can do better by applying the optimal integer aperture estimator to be presented there.

3.6 The Bayesian approach

The Bayesian approach to ambiguity resolution is fundamentally different from the approach using integer least-squares, which is generally used in practice but lacks sound

validation criteria. The fundamental difference is that in the Bayesian approach not only the vector of observations y is considered random, but the vector of unknown parameters x as well. The concept of Bayesian estimation was described in e.g. (Betti et al. 1993; Gundlich and Koch 2002; Gundlich 2002; Teunissen 2001b).

According to Bayes' theorem the *posterior* density $p(x|y)$ is proportional to the likelihood function $p(y|x)$ and the *prior* density $p(x)$:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \tag{3.92}$$

The Bayes estimate of the random parameter vector is defined as the conditional mean

$$\hat{x}_{\text{Bayes}} = E\{x|y\} = \int xp(x|y)dx \tag{3.93}$$

This estimate minimizes the discrepancy between x and $\hat{x}$ on the average, where $\hat{x}$ is a function of y. So, if $L(x, \hat{x})$ is the measure of loss (or discrepancy), this amounts to solving the minimization problem

$$\min_{\hat{x}} E\{L(x, \hat{x})|y\} = \min_{\hat{x}} \int L(x, \hat{x})p(x|y)dx \tag{3.94}$$

If $L(x, \hat{x}) = \|x - \hat{x}\|^2_Q$ it follows that the Bayes estimate $\hat{x}_{\text{Bayes}}$ is indeed the solution to this minimization problem.

In order to apply the Bayesian approach to ambiguity resolution, with the parameter vector $x = [b^T \; a^T]^T$, a and b are assumed to be independent with priors

$$\begin{cases} p(a) \propto \sum\limits_{z \in \mathbb{Z}^n} \delta(a - z) & \text{(pulsetrain)} \\ p(b) \propto constant \end{cases} \tag{3.95}$$

where δ is the Dirac function. From the orthogonal decomposition of equation (3.14) follows that the likelihood function

$$p(y|a, b) \propto \exp\{-\frac{1}{2}\left(\|\hat{a} - a\|^2_{Q_{\hat{a}}} + \|\hat{b}(a) - b\|^2_{Q_{\hat{b}|\hat{a}}}\right)\} \tag{3.96}$$

and the posterior density follows therefore as

$$p(a, b|y) \propto \exp\{-\frac{1}{2}\left(\|\hat{a} - a\|^2_{Q_{\hat{a}}} + \|\hat{b}(a) - b\|^2_{Q_{\hat{b}|\hat{a}}}\right)\} \sum_{z \in \mathbb{Z}^n} \delta(a - z) \tag{3.97}$$

The marginal posterior densities $p(a|y)$ and $p(b|y)$ follow from integrating this joint posterior density over the domains $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^p$ respectively. Note that the integration domain of a is not chosen as $\mathbb{Z}^n$ since the discrete nature of a is considered to be captured by assuming the prior to be a pulsetrain.

The marginal posterior PDFs are then obtained as

$$\begin{cases} p(a|y) = w_a(\hat{a}) \sum\limits_{z \in \mathbb{Z}^n} \delta(a - z) \\ p(b|y) = \sum\limits_{z \in \mathbb{Z}^n} p_{b|a}(b|a = z, y) w_z(\hat{a}) \end{cases} \tag{3.98}$$

with the weight function

$$w_z(\hat{a}) = \frac{\exp\{-\frac{1}{2}\|\hat{a}-z\|^2_{Q_{\hat{a}}}\}}{\sum\limits_{z\in\mathbb{Z}^n}\exp\{-\frac{1}{2}\|\hat{a}-z\|^2_{Q_{\hat{a}}}\}}, \quad z\in\mathbb{Z}^n \tag{3.99}$$

and conditional posterior

$$p_{b|a}(b|a,y) = \frac{1}{\sqrt{|Q_{\hat{b}|\hat{a}}|}(2\pi)^{\frac{p}{2}}}\exp\{-\frac{1}{2}\|\hat{b}(a)-b\|^2_{Q_{\hat{b}|\hat{a}}}\} \tag{3.100}$$

The Bayes baseline estimate can now be determined with equations (3.93) and (3.98) as

$$\hat{b}_{\text{Bayes}} = \int bp(b|y)db = \hat{b} - Q_{\hat{b}\hat{a}}Q_{\hat{a}}^{-1}(\hat{a} - \sum_{z\in\mathbb{Z}^n} zw_z(\hat{a})) \tag{3.101}$$

So, with the Bayesian approach the ambiguities are not resolved as integers. Instead a weighted sum over all possible ambiguities is used with proper weights derived from the likelihood function. A discussion on the corresponding confidence regions can be found in Gundlich and Koch (2002) and Gundlich (2002).

4 Best Integer Equivariant estimation

The integer least-squares estimator is the optimal choice from the class of admissible integer estimators as defined in definition 3.1.1, since it maximizes the success rate. However, it is not guaranteed that the fixed baseline estimator is always closer to the true but unknown baseline b than the float baseline estimator. Only if the success rate is very close to one, the fixed baseline estimator has a higher probability of being close to b than the float estimator, see section 3.4.2.

Since the success rate will not always be high, it might be interesting to have a baseline estimator that in some sense will always outperform the float estimator. In Teunissen (2003f) this estimator was introduced. It is the *best integer equivariant* (BIE) estimator and is based on a new class of estimators.

In this chapter, first the theory of integer equivariant estimation is introduced. This theory is then used to define the BIE estimators of the baseline parameters and ambiguities. The proofs were given in Teunissen (2003f). A more detailed description of the theory can be found in appendix C. Section 4.2 describes how the BIE estimators can be approximated. Finally, in section 4.3 the BIE estimator is compared with the float and fixed estimators.

4.1 The BIE estimator

The new class of estimators will have to fulfill two conditions. Firstly, condition (iii) of definition 3.1.1, which is the integer remove-restore property. Hence the name integer equivariant that is assigned to all estimators belonging to this class. The second condition is that if an arbitrary constant $\varsigma \in \mathbb{R}^p$ is added to the float baseline, the fixed baseline estimator should be shifted by the same amount. These conditions lead to the following definition of integer equivariant estimators.

Definition 4.1.1 *Integer equivariant estimators*

The estimator $\hat{\theta}_{\mathsf{IE}} = f_\theta(y)$ with $f_\theta : \mathbb{R}^m \mapsto \mathbb{R}$ is said to be an *integer equivariant* estimator of $\theta = l_a^T a + l_b^T b$ if

$$\begin{cases} f_\theta(y + Az) = f_\theta(y) + l_a^T z & \forall y \in \mathbb{R}^m, z \in \mathbb{Z}^n \\ f_\theta(y + B\varsigma) = f_\theta(y) + l_b^T \varsigma & \forall y \in \mathbb{R}^m, \varsigma \in \mathbb{R}^p \end{cases} \tag{4.1}$$

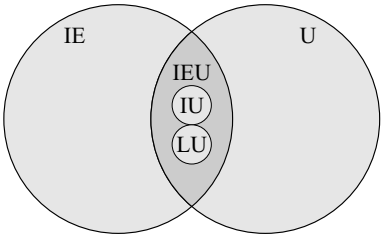

Figure 4.1: The set of relationships between the different classes of estimators: integer equivariant estimators (IE), unbiased integer equivariant estimators (IEU), unbiased integer estimators (IU), and linear unbiased estimators (LU).

By its definition, the class of IE estimators is larger than the class of admissible estimators defined in section 3.1, and thus all admissible estimators are also integer equivariant. Recall that the admissible integer estimators are unbiased, and therefore this class is referred to here as the class of integer unbiased (IU) estimators.

The class of IE estimators is also larger than the class of linear unbiased (LU) estimators. Let $f_\theta^T y$ be the linear estimator of $\theta = l_a^T a + l_b^T b$ for some $f_\theta \in \mathbb{R}^m$. This estimator is unbiased if $f_\theta^T Aa + f_\theta^T Bb = l_a^T a + l_b^T b$, $\forall a \in \mathbb{R}^n, b \in \mathbb{R}^m$ with $E\{y\} = Aa + Bb$. This gives:

$$\begin{cases} f_\theta(y + Aa) = f_\theta(y) + l_a^T a & \forall y \in \mathbb{R}^m, a \in \mathbb{R}^n \\ f_\theta(y + Bb) = f_\theta(y) + l_b^T b & \forall y \in \mathbb{R}^m, b \in \mathbb{R}^p \end{cases} \tag{4.2}$$

Comparing this result with (4.1) shows indeed that the condition of linear unbiasedness is more restrictive than that of integer equivariance, since the first condition must be fulfilled for all $a \in \mathbb{R}^n$. Hence, the class of linear unbiased estimators is a subset of the class of integer equivariant (IE) estimators, so that there must exist IE estimators that are unbiased. Summarizing, this gives the following relationships: $IEU = IE \cap U \neq \emptyset$, $LU \subset IEU$, and $IU \subset IEU$ (see figure 4.1).

The goal was to find an estimator that is 'best' in a certain sense. The criterion of 'best' that will be used is that the mean squared error (MSE) should be minimal. The reason for this choice is that the MSE is a well-known probabilistic criterion to measure the closeness of an estimator to its target value. Furthermore, the MSE-criterion is often used as a measure for the quality of the float solution itself. With the MSE-criterion the best integer equivariant (BIE) estimator, $\tilde{\theta}$, is defined as:

$$\tilde{\theta} = \arg \min_{f_\theta \in IE} E\{(f_\theta(y) - \theta)^2\} \tag{4.3}$$

Note that the minimization is taken over all IE functions that satisfy the condition of definition 4.1.1. The solution of this minimization problem is given by:

$$\tilde{\theta} = \frac{\sum\limits_{z \in \mathbb{Z}^n} \int\limits_{\mathbb{R}^p} (l_a^T z + l_b^T \zeta) f_y\left(y + A(a - z) + B(b - \zeta)\right) d\zeta}{\sum\limits_{z \in \mathbb{Z}^n} \int\limits_{\mathbb{R}^p} f_y\left(y + A(a - z) + B(b - \zeta)\right) d\zeta} \tag{4.4}$$

where $y \in \mathbb{R}^m$ with mean $E\{\underline{y}\} = Aa + Bb$ and PDF $f_{\underline{y}}(y)$, and $\tilde{\theta}$ is the BIE estimator of $\theta = l_a^T a + l_b^T b$.

The BIE estimator can also be written as:

$$\tilde{\theta} = l_a^T \tilde{a} + l_b^T \tilde{b} \tag{4.5}$$

with

$$\begin{cases} \tilde{a} = \sum\limits_{z \in \mathbb{Z}^n} z w_z(y), & \sum\limits_{z \in \mathbb{Z}^n} w_z(y) = 1 \\ \tilde{b} = \int\limits_{\mathbb{R}^p} \zeta w_\zeta(y) d\zeta, & \int\limits_{\mathbb{R}^p} w_\zeta(y) d\zeta = 1 \end{cases} \tag{4.6}$$

where the weighting functions $w_z(y)$ and $w_\zeta(y)$ follow from equation (4.4). This shows that the BIE estimator of the integer parameter vector a is also a weighted sum of all integer vectors in $\mathbb{Z}^n$, just like the admissible integer estimators of (3.6). However, in this case the weights are not binary (1 or 0); their values are determined by y and its PDF, so that the BIE weights are real-valued and nonzero for all integer vectors. This implies that the estimator $\tilde{a}$ is real-valued, instead of integer-valued. Furthermore, it implies that there is no need to apply a discrimination test.

In Teunissen (2003f) it was shown that the BIE estimator is unbiased and has minimum variance, i.e. it has better precision than the best linear unbiased estimator. This result applied to the baseline estimator gives:

$$\begin{cases} D\{\tilde{b}\} \leq D\{\check{b}\} \quad \text{and} \quad E\{\tilde{b}\} = E\{\check{b}\} = b \\ \\ D\{\tilde{b}\} \leq D\{\hat{b}\} \quad \text{and} \quad E\{\tilde{b}\} = E\{\hat{b}\} = b \end{cases} \tag{4.7}$$

Hence, the BIE baseline precision is always better than or as good as the precision of its float and fixed counterparts.

In GNSS applications it is assumed that the data are normally distributed as mentioned in section 3.1.3. It can be shown that in this case the BIE estimator is given by equation (4.5) with the following estimators for the ambiguities and the baseline parameters:

$$\begin{cases} \tilde{a} = \dfrac{\sum\limits_{z \in \mathbb{Z}^n} z \exp\{-\frac{1}{2}\|\hat{a}-z\|^2_{Q_{\hat{a}}}\}}{\sum\limits_{z \in \mathbb{Z}^n} \exp\{-\frac{1}{2}\|\hat{a}-z\|^2_{Q_{\hat{a}}}\}} = \sum\limits_{z \in \mathbb{Z}^n} z w_z(\hat{a}) \\ \\ \tilde{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \tilde{a}) \end{cases} \tag{4.8}$$

See appendix C.

Note that the formal expressions in (4.8) are identical to their Bayesian counterparts as presented in section 3.6, but that the distributional properties of the BIE estimator and its Bayesian counterpart of course differ. Since the functional form of the non-Bayesian estimator is identical to the Bayesian solution in the special case of normally distributed data, the theory of BIE estimation has provided the link with the Bayesian approach of ambiguity resolution.

4.2 Approximation of the BIE estimator

The BIE ambiguity estimator cannot be computed exactly because of the infinite sum in (4.8). If the infinite sum is replaced by a sum over a finite set of integers, Θ, this might result in an estimator that is **not** integer equivariant anymore. Therefore in this section it will be described how this finite integer set should be chosen, see (Teunissen 2004c).

4.2.1 The integer set

In order to determine a finite set of integers for the approximation of the BIE ambiguity estimator without spoiling the property of integer equivariance, this finite set Θ should not be chosen as a fixed set, but as a set of integers that depends on the float ambiguity vector $\hat{a}$. This can be attained by defining an ellipsoidal region:

$$C_0^\lambda = \left\{x \in \mathbb{R}^n \mid \|x\|^2_{Q_{\hat{a}}} < \lambda^2\right\} \tag{4.9}$$

so that $C_0^\lambda + z = C_z^\lambda = \left\{x \in \mathbb{R}^n \mid \|x - z\|^2_{Q_{\hat{a}}} < \lambda^2\right\}$. This set has the indicator function:

$$\omega_z^\lambda(x) = \begin{cases} 1 & \text{if } x \in C_z^\lambda \\ 0 & \text{otherwise} \end{cases}$$

Furthermore the following integer set is defined:

$$\Theta_x^\lambda = \left\{z \in \mathbb{Z}^n \mid \|x - z\|^2_{Q_{\hat{a}}} < \lambda^2\right\} \tag{4.10}$$

With this integer set the BIE estimator of (4.8) can be approximated as:

$$\begin{aligned} \tilde{a}^\lambda &= \frac{\sum\limits_{z\in\mathbb{Z}^n} z\omega_z^\lambda(\hat{a}) \exp\{-\frac{1}{2}\|\hat{a} - z\|^2_{Q_{\hat{a}}}\}}{\sum\limits_{z\in\mathbb{Z}^n} \omega_z^\lambda(\hat{a}) \exp\{-\frac{1}{2}\|\hat{a} - z\|^2_{Q_{\hat{a}}}\}} \\ &= \frac{\sum\limits_{z\in\Theta_{\hat{a}}^\lambda} z \exp\{-\frac{1}{2}\|\hat{a} - z\|^2_{Q_{\hat{a}}}\}}{\sum\limits_{z\in\Theta_{\hat{a}}^\lambda} \exp\{-\frac{1}{2}\|\hat{a} - z\|^2_{Q_{\hat{a}}}\}} = \sum_{z\in\Theta_{\hat{a}}^\lambda} z w_z^\lambda(\hat{a}) \end{aligned} \tag{4.11}$$

This estimator is still integer equivariant. Obviously, $\lim\limits_{\lambda\to\infty} \tilde{a}^\lambda = \tilde{a}$. The approximation $\tilde{a}^\lambda$ would become equal to the true BIE solution when the PDF of $\hat{a}$ would be given by the truncated normal distribution:

$$f_{\hat{a}}^\lambda(x) = \frac{\omega_a^\lambda(x) \exp\{-\frac{1}{2}\|x - a\|^2_{Q_{\hat{a}}}\}}{\int\limits_{\Theta_a^\lambda} \exp\{-\frac{1}{2}\|x - a\|^2_{Q_{\hat{a}}}\}} \tag{4.12}$$

with $\Theta_a^\lambda = \left\{x \in \mathbb{R}^n \mid \|x - a\|^2 < \lambda^2\right\}$.

Since the BIE estimator is a weighted sum over all integers, the integer set should be chosen such that the weights $w_u(\hat{a})$, $\forall u \notin \Theta_{\hat{a}}^\lambda$ are so small that $w_u(\hat{a})u \approx 0$. The

difference between the BIE estimator and its approximation is:

$$\begin{aligned}\tilde{a} - \tilde{a}^{\lambda} &= \sum_{z \in \mathbb{Z}^n} z w_z(\hat{a}) - \sum_{z \in \Theta_{\hat{a}}^{\lambda}} z w_z^{\lambda}(\hat{a}) \\ &= \sum_{z \in \Theta_{\hat{a}}^{\lambda}} z(w_z(\hat{a}) - w_z^{\lambda}(\hat{a})) + \sum_{z \in \mathbb{Z}^n \setminus \{\Theta_{\hat{a}}^{\lambda}\}} z w_z(\hat{a}) \end{aligned} \tag{4.13}$$

Note that $w_z^{\lambda}(\hat{a}) \geq w_z(\hat{a})$, since for the denominators of equations (4.11) and (4.8) it follows that

$$\sum_{z \in \Theta_{\hat{a}}^{\lambda}} \exp\{-\frac{1}{2}\|\hat{a} - z\|_{Q_{\hat{a}}}^2\} \leq \sum_{z \in \mathbb{Z}^n} \exp\{-\frac{1}{2}\|\hat{a} - z\|_{Q_{\hat{a}}}^2\}$$

Hence, the approximation error is not only the result of ignoring the last term on the right-hand side of equation (4.13), but also due to the different weights assigned to the integers in the set $\Theta_{\hat{a}}^{\lambda}$.

In order to get a good approximation, λ can be determined with:

$$P(\hat{a} \in \Theta_a^{\lambda}) = P(\|\hat{a} - a\|_{Q_{\hat{a}}}^2 < \lambda^2) = 1 - \alpha \tag{4.14}$$

since it is known that $\|\hat{a} - a\|_{Q_{\hat{a}}}^2$ has a central χ^2-distribution with n degrees of freedom.

4.2.2 Z-transformation

In section 3.1.4 it was explained that the Z-transformation makes the search for integers within an ellipsoidal region much more efficient. For the approximation of the BIE estimator also a set of integers within an ellipsoidal region must be determined. Therefore, BIE estimation should be applied to the Z-transformed float ambiguities, since the transformation is also admissible in this case:

$$\begin{aligned}\tilde{z}^{\lambda} &= \frac{\sum_{z \in \Theta_{\hat{z}}^{\lambda}} z \exp\{-\frac{1}{2}\|\hat{z} - z\|_{Q_{\hat{z}}}^2\}}{\sum_{z \in \Theta_{\hat{z}}^{\lambda}} \exp\{-\frac{1}{2}\|\hat{z} - z\|_{Q_{\hat{z}}}^2\}} \\ &= \frac{\sum_{z \in \Theta_{\hat{z}}^{\lambda}} Z^T Z^{-T} z \exp\{-\frac{1}{2}\|\hat{a} - Z^{-T} z\|_{Q_{\hat{a}}}^2\}}{\sum_{z \in \Theta_{\hat{z}}^{\lambda}} \exp\{-\frac{1}{2}\|\hat{a} - Z^{-T} z\|_{Q_{\hat{a}}}^2\}} \\ &= \frac{\sum_{z \in \Theta_{\hat{a}}^{\lambda}} z \exp\{-\frac{1}{2}\|\hat{a} - z\|_{Q_{\hat{a}}}^2\}}{\sum_{z \in \Theta_{\hat{a}}^{\lambda}} \exp\{-\frac{1}{2}\|\hat{a} - z\|_{Q_{\hat{a}}}^2\}} \qquad (4.15) \\ &= Z^T \tilde{a}^{\lambda} \qquad (4.16)\end{aligned}$$

4.3 Comparison of the float, fixed, and BIE estimators

Although it is known that the BIE estimator outperforms its float and fixed counterparts in terms of precision, it is not clear how large this difference will be under varying measurement scenarios.

It follows that in the limits of the precision the following is true:

$$\lim_{\sigma\to\infty} \tilde{a} = \hat{a} \quad \text{and} \quad \lim_{\sigma\to 0} \tilde{a} = \check{a} \tag{4.17}$$

with the vc-matrix factored as $Q_{\hat{a}} = \sigma^2 G_{\hat{a}}$. The first limit corresponds to an integer grid size which is very small in relation to the size and extend of the PDF of the float ambiguities. This implies that the summation over all integers will be approximately equal to the integration over the space of reals, and then:

$$\tilde{a} = \frac{\int\limits_{\mathbb{R}^n} x \exp\{-\frac{1}{2}\|\hat{a}-x\|^2_{Q_{\hat{a}}}\}dx}{\int\limits_{\mathbb{R}^n} \exp\{-\frac{1}{2}\|\hat{a}-x\|^2_{Q_{\hat{a}}}\}dx} = \frac{\hat{a}(2\pi)^{\frac{n}{2}}\sqrt{|Q_{\hat{a}}|}}{(2\pi)^{\frac{n}{2}}\sqrt{|Q_{\hat{a}}|}} = \hat{a} \tag{4.18}$$

This shows that if the precision is low the BIE estimator will approximate the float estimator. If, on the other hand, the PDF of the float ambiguities becomes very peaked the BIE estimator will approximate the fixed estimator. This can be shown as follows.

Let the ILS pull-in region of $z \in \mathbb{Z}^n$ be given as $S_z = \{x \in \mathbb{R}^n \mid \|x-z\|^2_{Q_{\hat{a}}} \leq \|x-u\|^2_{Q_{\hat{a}}},\ \forall u \in \mathbb{Z}^n\}$. Then S_z is independent of σ, and the weight $w_z(x)$ is equal to (see equation (4.8)):

$$\begin{aligned} w_z(x) &= \frac{\exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}}{\sum\limits_{u\in\mathbb{Z}^n}\exp\{-\frac{1}{2}\|x-u\|^2_{Q_{\hat{a}}}\}} \\ &= \frac{1}{1+\sum\limits_{u\in\mathbb{Z}^n\setminus\{z\}}\exp\{-\frac{1}{2\sigma^2}(\|x-u\|^2_{G_{\hat{a}}}-\|x-z\|^2_{G_{\hat{a}}})\}} \end{aligned} \tag{4.19}$$

It follows thus that $\lim\limits_{\sigma\to 0} w_z(x) = 1$ if $x \in S_z$, and since the sum of all weights must equal one $\lim\limits_{\sigma\to 0} w_u(x) = 0$, $\forall u \neq z$ if $x \in S_z$. So, $w_u(x)$ reduces to the indicator function of the ILS pull-in region in the limit $\sigma \to 0$.

Obviously, the limiting cases in equation (4.17) imply that also the following is true:

$$\lim_{\sigma\to\infty}\|\hat{a}-\tilde{a}\|^2_{Q_{\hat{a}}} = 0 \quad \text{and} \quad \lim_{\sigma\to 0}\|\hat{a}-\tilde{a}\|^2_{Q_{\hat{a}}} = \|\hat{a}-\check{a}\|^2_{Q_{\hat{a}}} \tag{4.20}$$

It is now interesting to know how the BIE estimator performs in the intermediate cases compared to the float and fixed estimators. In this section the three estimators are therefore compared numerically. This also provides the possibility to compare their distributional properties. The results were presented in Verhagen and Teunissen (2003).

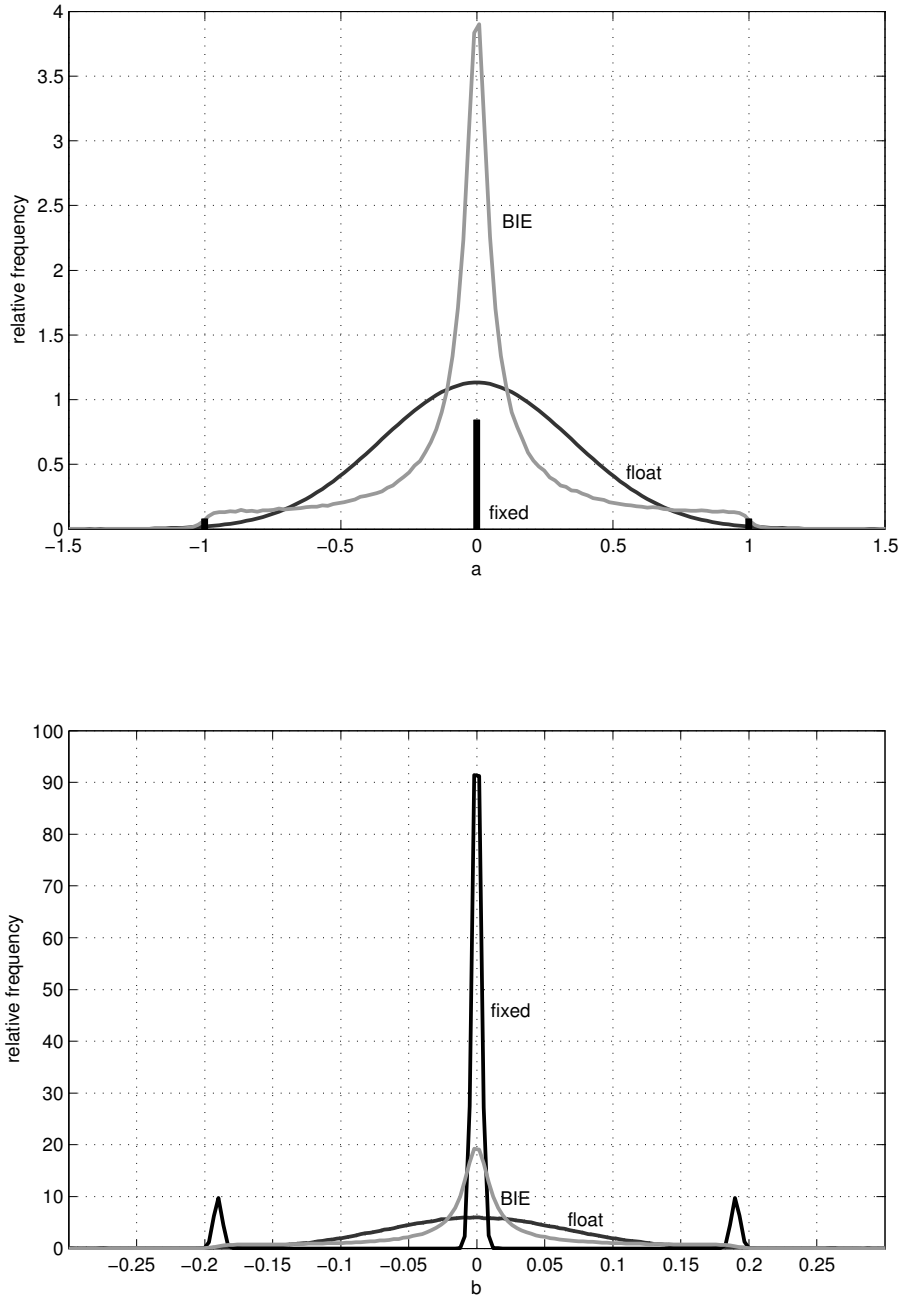

Figure 4.2: Examples of distributions of the different estimators. Top: $\hat{a}$ (solid), $\check{a}$ (bars), $\tilde{a}$ (dashed); Bottom: $\hat{b}$ (solid), $\check{b}$ (solid, multi-modal), $\tilde{b}$ (dashed).

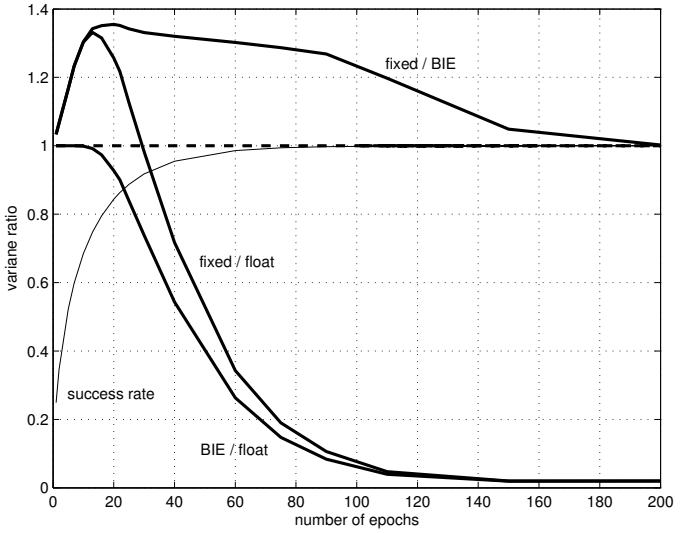

Figure 4.3: Variance ratios of: BIE and float estimator; BIE and fixed estimator; fixed and float estimator. Success rate as function of the number of epochs is also shown.

4.3.1 The 1-D case

Monte Carlo simulations were carried out to generate 500,000 samples of the float range and ambiguity, using the geometry-free single frequency GPS model for k epochs, with vc-matrix:

$$\begin{pmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}_k} \\ Q_{\hat{b}_k\hat{a}} & Q_{\hat{b}_k} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_p^2}{k\lambda^2}(1+\varepsilon) & -\frac{\sigma_p^2}{k\lambda^2} \\ -\frac{\sigma_p^2}{k\lambda^2} & \frac{\sigma_p^2}{1+\varepsilon}(\frac{1}{k}+\varepsilon) \end{pmatrix}$$

with λ the wavelength of the carrier; σ_p^2 and σ_ϕ^2 are the variances of the DD code and phase observations respectively, and $\varepsilon = \sigma_\phi^2/\sigma_p^2$. For all simulations, the standard deviations were chosen as $\sigma_p = 30$ cm and $\sigma_\phi = 3$ mm. The number of epochs was varied. Note that in the one-dimensional case the fixed ambiguity estimator is obtained by simply rounding the float estimator to the nearest integer.

Figure 4.2 shows the parameter distributions of all three estimators for $k = 20$, based on the simulation results. Note that the multi-modality of the distribution of the BIE range estimator is less pronounced than that of the fixed range estimator. For smaller k the distribution of the BIE ambiguity and range estimator would resemble the normal PDF of the float estimators. For larger k, and thus higher precision, the distribution of the BIE estimators would more and more resemble those of the fixed estimators.

From equation (4.7) follows that the BIE baseline estimator has smallest variance, but in the limits of the precision the variance will become equal to the variance of the float and fixed estimator. This is illustrated in figure 4.3, where the variance ratio of the different estimators is shown as function of k. Also the success rate is shown. Indeed, for small k the variance of the BIE and the float estimator are equal to each other (ratio equals

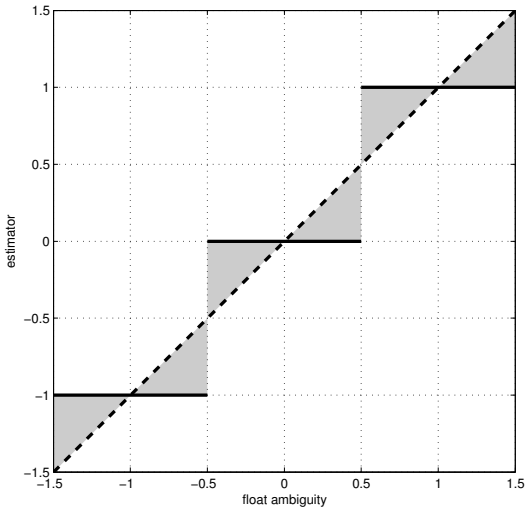

Figure 4.4: Ambiguity estimators as function of the float ambiguity. Solid lines: fixed estimator; Dashed line: float estimator; the BIE estimator will always fall in the grey area.

one), and smaller than the variance of the fixed estimator. On the other hand, for large k the variances of the BIE and fixed estimator become equal to each other, and smaller than that of the float estimator. Only after 30 epochs, when the success rate is larger than 0.9, the variance of the fixed baseline estimator becomes lower than the variance of the float estimator.

It can be shown that in the one-dimensional case $|\hat{a} - \tilde{a}| \leq |\hat{a} - \breve{a}|$, $\forall \hat{a}, \sigma^2$. First note that:

$$|\hat{a} - \tilde{a}| = 0 \quad \text{if} \quad \hat{a} = \frac{1}{2}z, z \in \mathbb{Z} \tag{4.21}$$

Furthermore, $\tilde{a}$ is monotonically increasing with $\hat{a}$, since:

$$\begin{aligned}
\frac{d\tilde{a}}{d\hat{a}} &= \frac{1}{\Gamma^2}\left[\sum_{z\in\mathbb{Z}} z(z-x)\sigma^{-2}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}\sum_{z\in\mathbb{Z}}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}-\right.\\
&\qquad\left.-\sum_{z\in\mathbb{Z}} z\sigma^{-2}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}\sum_{z\in\mathbb{Z}}(z-x)\sigma^{-2}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}\right]\\
&= \frac{1}{\sigma^2\Gamma^2}\sum_{u\in\mathbb{Z}}\sum_{z\in\mathbb{Z}}\left[u(u-x)\exp\{-\frac{1}{2\sigma^2}(x-u)^2\}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}\right.\\
&\qquad\left.-u(z-x)\exp\{-\frac{1}{2\sigma^2}(x-u)^2\}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}\right]\\
&= \frac{1}{\sigma^2\Gamma^2}\sum_{u\in\mathbb{Z}}\sum_{z\in\mathbb{Z}}\underbrace{(u^2-uz)\exp\{-\frac{1}{2\sigma^2}(x-u)^2\}\exp\{-\frac{1}{2\sigma^2}(x-z)^2\}}_{F(u,z)}\\
&\geq 0
\end{aligned} \tag{4.22}$$

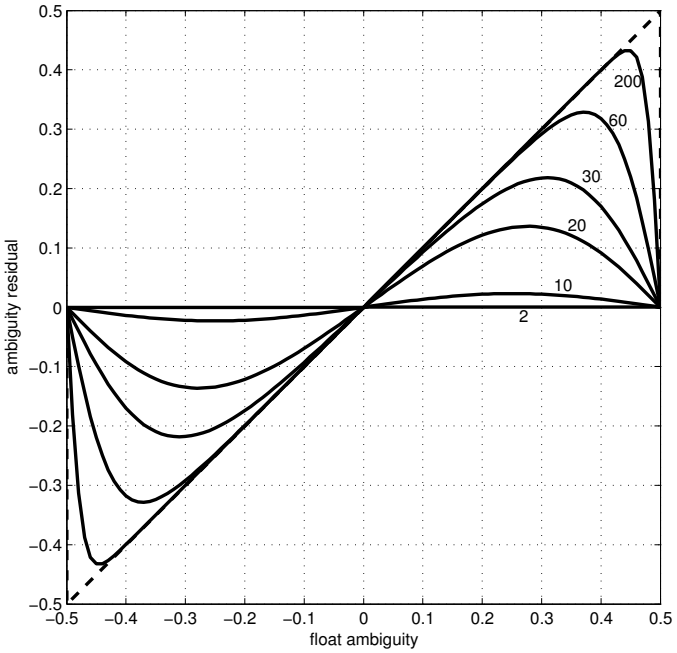

Figure 4.5: Ambiguity residuals for fixed (dashed) and BIE (solid) estimators for different number of epochs.

with $\Gamma^2 = \sum_{z\in\mathbb{Z}} \exp\{-\frac{1}{2\sigma^2}(x-z)^2\}$. The inequality follows from:

$$\begin{cases} F(z,z) = 0, & \forall z \in \mathbb{Z} \\ F(u,z) + F(z,u) \geq 0, & \forall u, z \in \mathbb{Z}, u \neq z \end{cases} \tag{4.23}$$

The results in equations (4.17), (4.21) and (4.22) imply that $\tilde{a}$ will always fall in the grey region in figure 4.4, and thus that indeed $|\hat{a} - \tilde{a}| \leq |\hat{a} - \check{a}|$. This means that $\tilde{a}$ will always lie in-between $\hat{a}$ and $\check{a}$, and thus that the BIE and the fixed estimator are pulled in the same direction in the one-dimensional case. This is shown in figure 4.5 for different precisions (i.e. for different k). The ambiguity residuals are defined as $\check{\epsilon} = \hat{a} - \check{a}$ and $\tilde{\epsilon} = \hat{a} - \tilde{a}$.

Figure 4.6 shows the probability that the baseline estimators will be within a certain interval 2ε that is centered at the true baseline b, again for different precisions. It can be seen that for high success rates indeed relation $P(\check{b} \in E_b) \geq P(\hat{b} \in E_b)$ is true, with E_b a convex region centered at b, cf. (Teunissen 2003c), which means that the probability that the fixed baseline will be closer to the true but unknown baseline is larger than that of the float baseline. However, for lower success rates some probability mass for $\check{b}$ can be located far from b because of the multi-modal distribution. Ideally, the probability should be high for small ε and reach its maximum as soon as possible. For lower success rates, the float and fixed estimators will only fulfill one of these conditions. The probability for the BIE estimator always falls more or less in-between those of the float and fixed estimators.

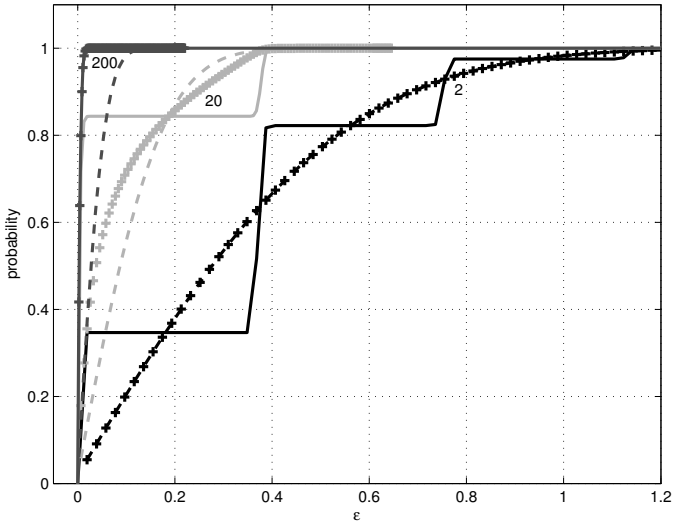

Figure 4.6: Probabilities $P(|\hat{b}-b| \leq \varepsilon)$ (dashed), $P(|\check{b}-b| \leq \varepsilon)$ (solid), $P(|\tilde{b}-b| \leq \varepsilon)$ (+-signs) for different number of epochs.

The probabilities shown in figure 4.6 are determined by counting the number of solutions that fall within a certain interval, but it is also interesting to compare the estimators on a sample by sample basis. In order to do so, one could determine for each sample which of the three estimators is closest to the true b, and then count for each estimator how often it was better than the other estimators.

In table 4.1 the probabilities

$$P_1 = P(|\check{b}-b| \leq |\hat{b}-b|), \quad P_2 = P(|\tilde{b}-b| \leq |\hat{b}-b|), \text{ and } \quad P_3 = P(|\tilde{b}-b| \leq |\check{b}-b|)$$

are given for different number of epochs. It follows that the probability that $\tilde{b}$ is better than the corresponding $\hat{b}$ is larger or equal to the probability that $\check{b}$ is better than $\hat{b}$. That is because the ambiguity residuals that are used to compute the fixed and BIE baseline estimator, see equations (3.4) and (4.8), have the same sign, and $|\hat{a}-\tilde{a}| \leq |\hat{a}-\check{a}|$ as was shown in figure 4.5. If the float baseline solution is already close to the true solution, it is possible that $\tilde{b}$ is closer to b, but that $\check{b}$ is pulled 'over' the true solution such that $\check{b}-b$ has the opposite sign as $\hat{b}-b$ and $|\check{b}-b| > |\hat{b}-b|$.

From the relationships between the BIE and the fixed ambiguity residuals it follows that if $\check{b}$ is better than $\hat{b}$, then also $\tilde{b}$ will be better than $\hat{b}$. In the case of very low precision $\tilde{a} \approx \hat{a}$ and $\tilde{b} \approx \hat{b}$, so that P_2 is approximately equal to one – in the example this is the case for $k = 1$ and $k = 2$. If the precision is high, $\tilde{a} \approx \check{a}$ and as a result P_3 becomes very high.

Note that these results do not necessarily hold true for the higher dimensional case ($n > 1$).

Table 4.1: Probabilities $P_1 = P(|\check{b} - b| \leq |\hat{b} - b|)$, $P_2 = P(|\tilde{b} - b| \leq |\hat{b} - b|)$, and $P_3 = P(|\tilde{b} - b| \leq |\check{b} - b|)$.

k	$P(\check{a} = a)$	P_1	P_2	P_3
1	0.249	0.557	1.000	0.443
2	0.346	0.583	1.000	0.417
5	0.522	0.635	0.642	0.365
10	0.685	0.707	0.716	0.296
20	0.844	0.820	0.826	0.240
30	0.918	0.885	0.887	0.301
60	0.986	0.937	0.937	0.458
75	0.994	0.939	0.939	0.480
110	0.999	0.933	0.933	0.552
200	1.000	0.913	0.913	0.982

4.3.2 The 2-D case

A similar approach was followed to generate samples for the geometry-free dual-frequency GPS model for one satellite-pair. The double difference standard deviations were chosen such that success rates between 0.3 and 0.99 were obtained, see table 4.2.

Figure 4.7 shows an example of the BIE and fixed ambiguity estimates that correspond to certain float ambiguity vectors ($\sigma_p = 0.8$m, $\sigma_\phi = 8$mm). It can be seen that the BIE estimator is pulled in approximately the same direction as the ILS estimator, but the ambiguity residuals are always much smaller for this example. If the float solution falls close to an integer grid point, the BIE estimator is pulled in that direction. On the other hand, if the float solution falls close to the boundary of a pull-in region, which means that the distance to at least two integer grid points is approximately the same, the ambiguity residual is small and the BIE estimator is pulled in another direction.

Figure 4.8 shows a scatter plot of simulated float and corresponding BIE ambiguity estimates ($\sigma_p = 0.6$m, $\sigma_\phi = 6$mm). For higher precisions, the 'star'-shape of the distribution of the BIE ambiguity estimator becomes even stronger; for lower precisions the distribution becomes more ellipse-shaped like that of the float estimator.

Table 4.2 shows the probabilities

$$P_1 = P(|\check{b}-b| \leq |\hat{b}-b|), \quad P_2 = P(|\tilde{b}-b| \leq |\hat{b}-b|), \text{ and } \quad P_3 = P(|\tilde{b}-b| \leq |\check{b}-b|)$$

This table only shows the probabilities that the estimators will perform better than or equal to the other estimators. However, this does not say how much better, or how much worse in the instances that they do not perform better. Therefore, one should also consider the probabilities as shown in figure 4.9. From the table it follows namely that $P(|\check{b} - b| \leq |\hat{b} - b|)$ is larger than $P(|\tilde{b} - b| \leq |\hat{b} - b|)$. On the other hand, figure 4.9 shows that if $\check{b}$ is not close to b, it will immediately be far from the true solution (because of the step-wise function), whereas for the BIE estimator that is not necessarily the case.

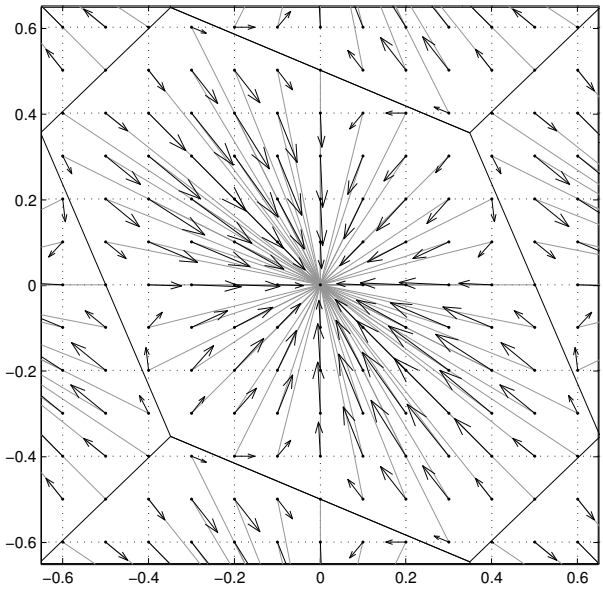

Figure 4.7: For different float ambiguity vectors the corresponding BIE and fixed estimator are determined. The arrows point from float to corresponding BIE estimator.

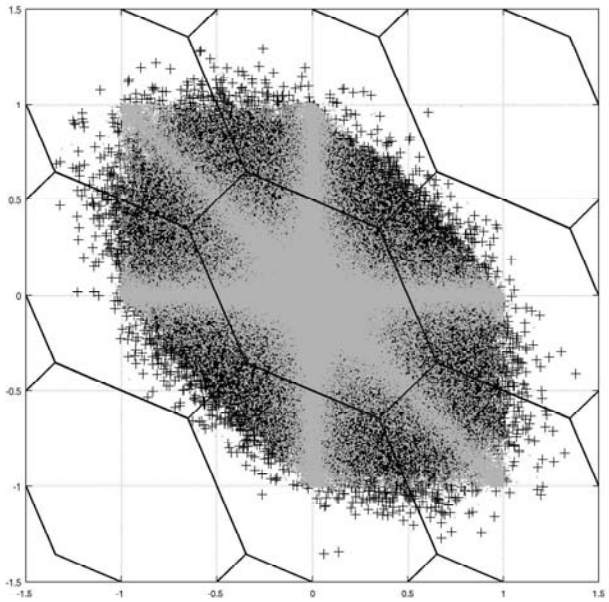

Figure 4.8: Scatter plot of float ambiguities (black) and BIE ambiguity estimates (grey).

Table 4.2: Probabilities $P_s = P(\check{a} = a)$, $P_1 = P(|\check{b}-b| \leq |\hat{b}-b|)$, $P_2 = P(|\tilde{b}-b| \leq |\hat{b}-b|)$, and $P_3 = P(|\tilde{b} - b| \leq |\check{b} - b|)$. Double difference code and phase standard deviations respectively: $\sigma_p = \sigma$m and $\sigma_\phi = \sigma$cm.

σ	P_s	P_1	P_2	P_3
1.4	0.312	0.526	0.521	0.474
0.8	0.675	0.724	0.708	0.283
0.6	0.859	0.875	0.859	0.205
0.4	0.985	0.983	0.981	0.434

Table 4.3: Mean and maximum approximation errors in $\|\hat{z}-\tilde{z}\|^2_{Q_{\hat{z}}}$. N_z equals the minimum and maximum number of integers used.

	α	mean	maximum	N_z
06_01	10^{-16}			753-850
	10^{-12}	10^{-11}	10^{-11}	347-407
	10^{-8}	10^{-7}	10^{-6}	112-155
	10^{-6}	10^{-6}	10^{-5}	53-87
06_02	10^{-16}			4730-4922
	10^{-12}	10^{-12}	10^{-12}	2203-2379
	10^{-8}	10^{-7}	10^{-7}	785-879
	10^{-6}	10^{-6}	10^{-6}	357-442
10_01	10^{-16}			274-345
	10^{-12}	10^{-11}	10^{-8}	80-121
	10^{-8}	10^{-7}	10^{-4}	11-31
	10^{-6}	10^{-5}	10^{-2}	2-14
10_03	10^{-16}			44323-45440
	10^{-12}	10^{-11}	10^{-10}	14353-15048
	10^{-8}	10^{-7}	10^{-6}	2959-3330
	10^{-6}	10^{-5}	10^{-4}	1007-1215

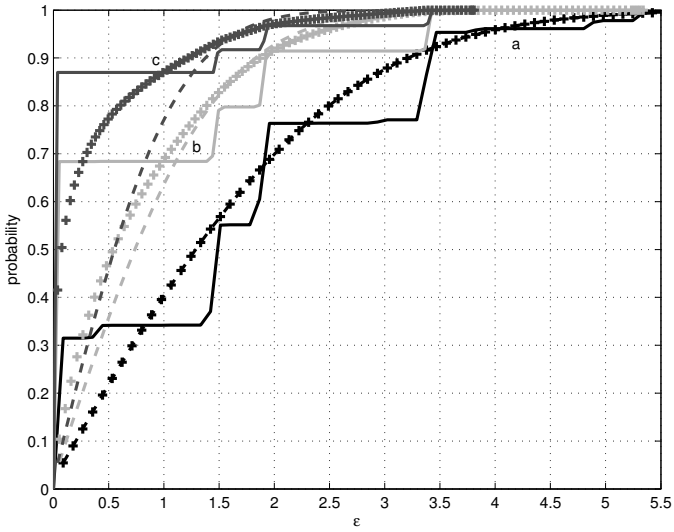

Figure 4.9: Probabilities $P(|\hat{b}-b| \leq \varepsilon]$ (dashed), $P(|\check{b}-b| \leq \varepsilon]$ (solid), $P(|\tilde{b}-b| \leq \varepsilon]$ (+-signs). a) $\sigma_p = 1.4$m, $\sigma_\phi = 14$mm; b) $\sigma_p = 0.8$m, $\sigma_\phi = 8$mm; c) $\sigma_p = 0.6$m, $\sigma_\phi = 6$mm.

4.3.3 The geometry-based case

Several geometry-based models were set up and corresponding float samples were generated. A description of the models is given in appendix B. Only the Z-transformed vc-matrices, $Q_{\hat{z}\_nn\_xx}$, are used.

First, it is investigated how the integer set Θ_x^λ must be chosen in order to get a good approximation for the BIE estimators. For that purpose, 10,000 samples of float ambiguities were generated and the integer set was determined using (4.10) and (4.14) for different values of α. Then the corresponding $\tilde{z}^\alpha$ were computed. Table 4.3 shows the difference $\|\hat{z}-\tilde{z}^\alpha\|^2_{Q_{\hat{z}}} - \|\hat{z}-\tilde{z}^{\alpha_{\min}}\|^2_{Q_{\hat{z}}}$ with $\alpha_{\min} = 10^{-16}$. Also shown are the number of integer vectors in the integer set. It follows that for all examples the solution converges to the true BIE estimator for $\alpha < 10^{-8}$.

The BIE estimators are compared based on the squared norms of the ambiguity residuals, and based on the probability that the baseline estimators are within a certain distance from the true baseline:

$$P(\|\dot{b}-b\|^2_{Q_{\dot{b}}} \leq \varepsilon) \tag{4.24}$$

where $\dot{b}$ is either the float, fixed, or BIE baseline estimator. The vc-matrices $Q_{\check{b}}$ and $Q_{\tilde{b}}$ are determined from the corresponding estimates. Since a user will be mainly interested in the baseline coordinates, also the following probabilities were analyzed:

$$P(\|\dot{b}_x - b_x\| \leq \varepsilon) \tag{4.25}$$

where b_x refers only to the three baseline coordinates. So, the distance to the true

Table 4.4: Probabilities that float, fixed or BIE baseline estimator is best.
$P_1(\dot{b}) = P(\|\dot{b} - b\|_{Q_{\dot{b}}} \leq \|\hat{b} - b\|_{Q_{\dot{b}}})$; $P_2 = P(\|\tilde{b} - b\|_{Q_{\tilde{b}}} \leq \|\check{b} - b\|_{Q_{\check{b}}})$;
$P_3(\dot{b}) = P(\|\dot{b} - b\|_{Q_{\dot{b}}} \leq \|\hat{b} - b\|_{Q_{\hat{b}}}, \|\dot{b} - b\|_{Q_{\dot{b}}} \leq \|\ddot{b} - b\|_{Q_{\ddot{b}}})$;
with $\ddot{b}$ the fixed or BIE estimator but $\ddot{b} \neq \dot{b}$.

	$P_1(\check{b})$	$P_1(\tilde{b})$	P_2	$P_3(\check{b})$	$P_3(\tilde{b})$
06_01	0.824	0.786	0.149	0.796	0.037
06_02	0.547	0.561	0.402	0.531	0.111
10_01	0.863	0.836	0.021	0.851	0.012
10_03	0.749	0.792	0.333	0.620	0.225

Table 4.5: Probabilities that float, fixed or BIE baseline estimator is best.
$P_1(\dot{b}_x) = P(\|\dot{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|)$; $P_2 = P(\|\tilde{b}_x - b\| \leq \|\check{b}_x - b_x\|)$;
$P_3(\dot{b}_x) = P(\|\dot{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|, \|\dot{b}_x - b_x\| \leq \|\ddot{b}_x - b_x\|)$;
with $\ddot{b}$ the fixed or BIE estimator but $\ddot{b} \neq \dot{b}$.

	$P_1(\check{b}_x)$	$P_1(\tilde{b}_x)$	P_2	$P_3(\check{b}_x)$	$P_3(\tilde{b}_x)$
06_01	0.862	0.852	0.248	0.746	0.147
06_02	0.603	0.612	0.417	0.566	0.152
10_01	0.996	0.996	0.482	0.518	0.479
10_03	0.885	0.909	0.430	0.563	0.363

position is considered here.

The results are shown in figures 4.10 and 4.11. The ambiguity residuals, $\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}} = \|\hat{z} - \tilde{z}\|^2_{Q_{\hat{z}}}$ and $\|\check{\epsilon}\|^2_{Q_{\hat{z}}} = \|\hat{z} - \check{z}\|^2_{Q_{\hat{z}}}$, are plotted for all samples, ordered by increasing $\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}}$. It can be seen that BIE ambiguity residuals are always smaller or equal than the ILS ambiguity residuals. Furthermore, the results can be explained as follows (compare with figures 4.5 and 4.7 for the 1-D and 2-D case respectively):

- $\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}}$ is small in two cases:

  1. $\hat{z}$ is close to an integer; then $\|\check{\epsilon}\|^2_{Q_{\hat{z}}}$ will also be small;
  2. $\hat{z}$ is close to the boundary of a pull-in region; then $\|\check{\epsilon}\|^2_{Q_{\hat{z}}}$ will be large. In this case $\|\check{\epsilon}\|^2_{Q_{\hat{z}}}$ can take a large range of values, since the boundary of the pull-in region is not ellipsoidal.

- $\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}}$ is large if $\hat{z}$ takes a value halfway an integer and the boundary of the pull-in region. In that case $\|\check{\epsilon}\|^2_{Q_{\hat{z}}}$ is relatively small.

- $\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}}$ and $\|\check{\epsilon}\|^2_{Q_{\hat{z}}}$ are approximately equal if $\hat{z}$ is relatively close to an integer. If the precision is poor the difference becomes larger.

Note that since $\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}} \leq \|\check{\epsilon}\|^2_{Q_{\hat{z}}}$ also the test statistic in equation (3.76) is smaller for the BIE estimator:

$$\frac{\|\tilde{\epsilon}\|^2_{Q_{\hat{z}}}}{n\hat{\sigma}^2} \leq \frac{\|\check{\epsilon}\|^2_{Q_{\hat{z}}}}{n\hat{\sigma}^2} \tag{4.26}$$

Therefore the BIE ambiguity estimator is more likely to be accepted than the fixed estimator. Moreover, there is no need for a discrimination test as there is only one candidate, which is obtained by weighting all integer vectors.

For the baseline probabilities similar conclusions can be drawn as for the 2-D case. For small ε the probability for the fixed baseline is generally much higher. Depending on the precision, however, the probability for the BIE estimator is sometimes comparable and increases faster for increasing ε, so that the probability of a large baseline error is smaller.

From the probabilities in tables 4.4 and 4.5 it follows that there is always a higher probability that the fixed estimator performs better than the other two estimators. The probabilities that the fixed and BIE estimator perform better than the float estimator are comparable.

4.4 Summary

If the ILS success rate is not close to one, the fixed ambiguity estimator may not be considered deterministic and the integer validation procedures that are currently available should not be used. So, in that case the ambiguities cannot be fixed reliably. But not fixing means that no use is made of the information that the ambiguities should be integer-valued. Therefore, the BIE estimator could be used as an alternative, since it makes use of the information that ambiguities are integer-valued although not in an optimal way, and it is defined such that it is best in the sense that the MSEs are minimized. As such, the BIE estimator can be considered as a compromise between the float and fixed estimator, since it approximates these estimators in the extreme cases of very low precision and very high precision respectively. This is also what follows from the examples shown in this chapter.

Disadvantages of BIE estimation are that it does still not provide an exact probabilistic evaluation of the final solution, and that it is computationally complex. Moreover, the results in this chapter indicate that only in a limited number of cases the BIE estimator performs obviously better than either the float or the fixed estimator.

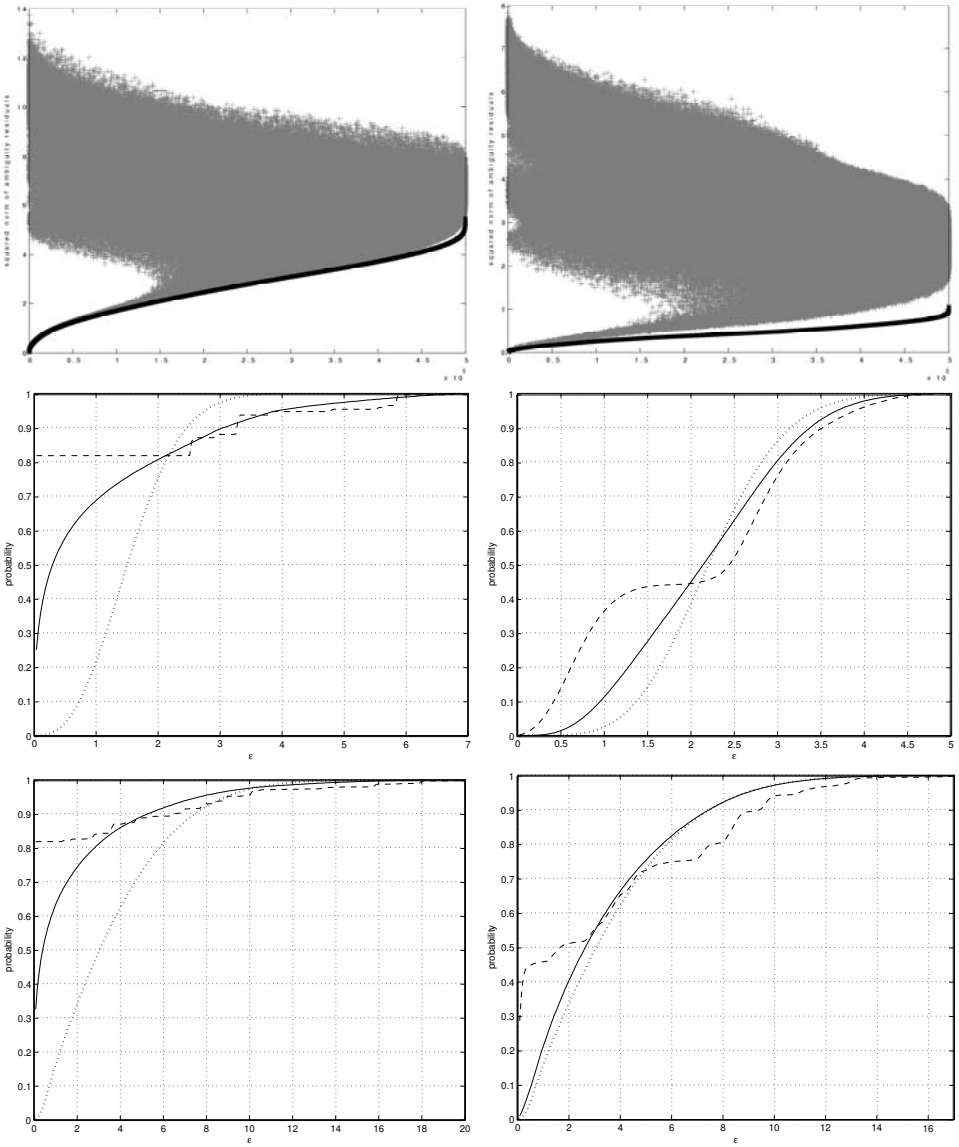

Figure 4.10: Example 06_01 (left) and 06_02 (right). Top: squared norm of ambiguity residuals for fixed estimator (grey +-signs) and BIE estimator (black x-sings). Center: $P(\|\dot{b} - b\|^2_{Q_{\dot{b}}} \leq \varepsilon)$. Bottom: $P(\|\dot{b}_x - b_x\|^2 \leq \varepsilon)$. Dotted: $\dot{b} = \hat{b}$; Dashed: $\dot{b} = \check{b}$; Solid: $\dot{b} = \tilde{b}$.

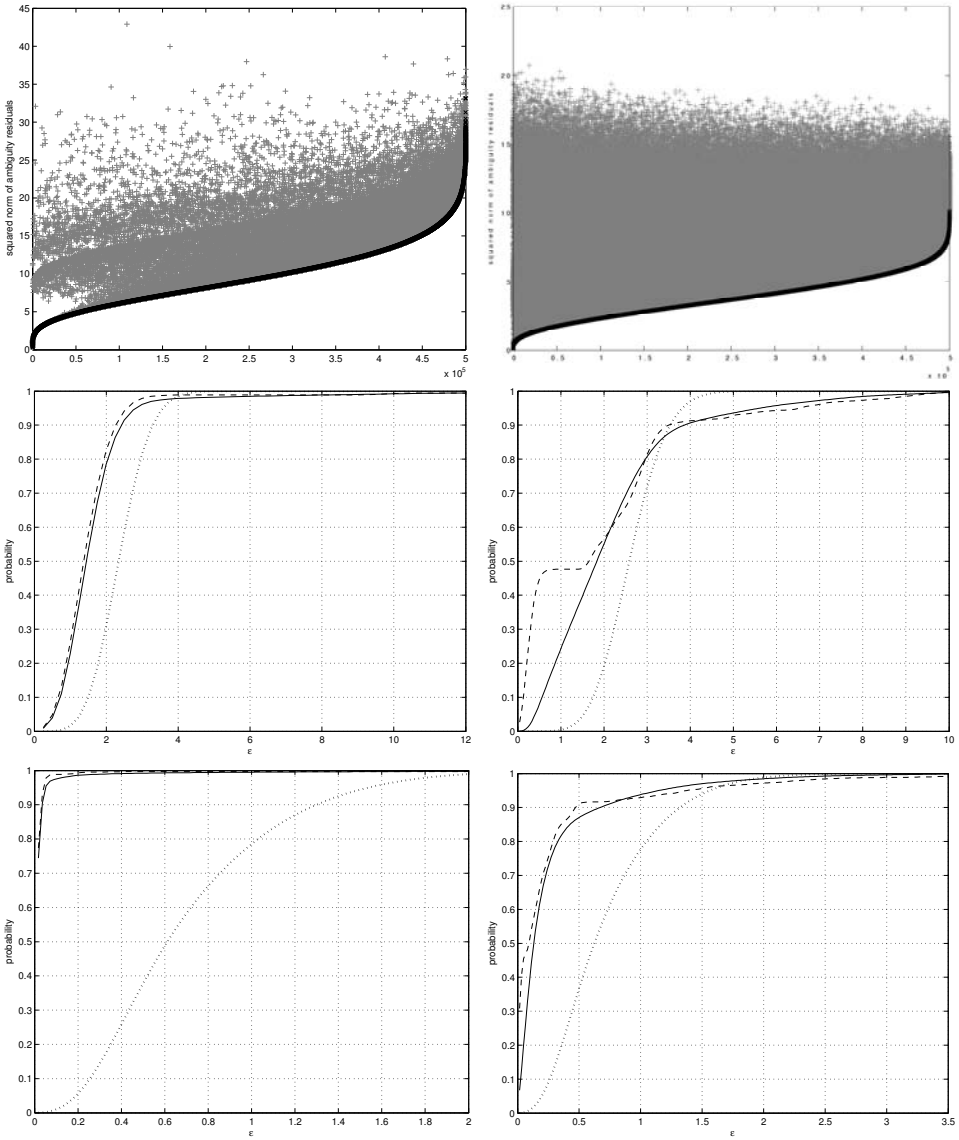

Figure 4.11: Example 10_01 (left) and 10_03 (right). Top: squared norm of ambiguity residuals for fixed estimator (grey +-signs) and BIE estimator (black x-sings). Center: $P(\|\dot{b} - b\|^2_{Q_{\hat{b}}} \leq \varepsilon)$. Bottom: $P(\|\dot{b}_x - b_x\|^2 \leq \varepsilon)$. Dotted: $\dot{b} = \hat{b}$; Dashed: $\dot{b} = \check{b}$; Solid: $\dot{b} = \tilde{b}$.

5 Integer Aperture estimation

In section 3.2 the success rate was introduced. This probability is a valuable measure for deciding whether one can have enough confidence in the fixed solution. As soon as the success rate is close enough to one, a user can fix the ambiguities, so that the precise carrier phase measurements start to contribute to the solution. However, when the success rate is considered too low with respect to a given threshold, it will not yet be possible to benefit from the precise phase observations. This also means that there is a probability that the user unnecessarily sticks to the float solution just because the fail rate is too high. In practice, not only the success rate is considered to decide on acceptance or rejection of the integer solution. The integer solution is also validated using one of the methods described in section 3.5. Since no sound validation criterion exists yet, in this chapter a new class of integer estimators is defined, the class of integer aperture estimators. Within this class estimators can be defined such that the acceptance region of the fixed solution is determined by the probabilities of success and failure.

The new class of integer estimators is defined in section 5.1. Then different examples of integer aperture estimators are presented. First, the ellipsoidal aperture estimator in section 5.2, for which the aperture space is built up of ellipsoidal regions. Then in section 5.3 it is shown that integer estimation in combination with the ratio test, difference test, or the projector test belongs to the class of integer aperture estimators. When the aperture space is chosen as a scaled pull-in region, the integer aperture bootstrapping and least-squares estimators can be defined. This will be shown in section 5.4. Finally, the penalized integer aperture estimator and optimal integer aperture estimator are presented in sections 5.5 and 5.6. Section 5.7 deals with implementation aspects. A comparison of all the different estimators is given in section 5.8, and examples in order to show the benefits of the new approach are given in section 5.9.

In order to illustrate the principle of IA estimation, throughout this chapter a two-dimensional (2-D) example will be used. Simulations were carried out to generate 500,000 samples of the float range and ambiguities, see appendix B, for the following vc-matrix:

$$Q = \begin{pmatrix} Q_{\hat{b}} & Q_{\hat{b}\hat{a}} \\ Q_{\hat{a}\hat{b}} & Q_{\hat{a}} \end{pmatrix} = \begin{pmatrix} 0.1800 & -0.1106 & 0.0983 \\ -0.1106 & 0.0865 & -0.0364 \\ 0.0983 & -0.0364 & 0.0847 \end{pmatrix}$$

Note that $Q_{\hat{a}} = Q_{\hat{z}\_02\_01}$.

5.1 Integer Aperture estimation

The problem described above is related to the definition of admissible integer estimators, which only distinguishes two situations: *success* if the float ambiguity falls inside the pull-in region S_a, and *failure* otherwise. In practice, a user will decide not to use the fixed solution when the probability of failure is too high. This gives rise to the thought that it might be interesting to consider not only the two above-mentioned situations, but also a third, namely *undecided*. Actually, with combined integer estimation and validation using a discrimination test, these three situations are also distinguished, but the discrimination tests can only be used in the case of high precision. Therefore a new class of integer estimators is defined that allows for the three situations mentioned above. Then integer estimators within this class can be chosen which somehow regulate the probability of each of those situations. This can be accomplished by considering only conditions (ii) and (iii) of definition 3.1.1, stating that the pull-in regions must be disjunct and translational invariant.

5.1.1 Class of IA estimators

The new class of ambiguity estimators was introduced in Teunissen (2003d;2003g), and is called the class of Integer Aperture (IA) estimators. It is defined as:

$$\begin{aligned}
&(i) \quad \bigcup_{z\in\mathbb{Z}^n} \Omega_z = \bigcup_{z\in\mathbb{Z}^n} (\Omega \cap S_z) = \Omega \cap (\bigcup_{z\in\mathbb{Z}^n} S_z) = \Omega \cap \mathbb{R}^n = \Omega \\
&(ii) \quad \Omega_u \cap \Omega_z = (\Omega \cap \Omega_u) \cap (\Omega \cap \Omega_z) = \Omega \cap (S_u \cap S_z) = \emptyset, \quad u, z \in \mathbb{Z}^n, u \neq z \\
&(iii) \quad \Omega_0 + z = (\Omega \cap S_0) + z = (\Omega + z) \cap (S_0 + z) = \Omega \cap S_z = \Omega_z, \quad \forall z \in \mathbb{Z}^n
\end{aligned} \tag{5.1}$$

$\Omega \subset \mathbb{R}^n$ is called the aperture space. From (i) follows that this space is built up of the Ω_z, which will be referred to as aperture pull-in regions. Conditions (ii) and (iii) state that these aperture pull-in regions must be disjunct and translational invariant. Note that the definition (5.1) is very similar to the definition of the admissible integer estimators (3.1.1), but that $\mathbb{R}^n$ is replaced by Ω.

The integer aperture estimator, $\bar{a}$, is now given by:

$$\bar{a} = \sum_{z\in\mathbb{Z}^n} z\omega_z(\hat{a}) + \hat{a}(1 - \sum_{z\in\mathbb{Z}^n} \omega_z(\hat{a})) \tag{5.2}$$

with the indicator function $\omega_z(x)$ defined as:

$$\omega_z(x) = \begin{cases} 1 & \text{if } x \in \Omega_z \\ 0 & \text{otherwise} \end{cases} \tag{5.3}$$

So, when $\hat{a} \in \Omega$ the ambiguity will be fixed using one of the admissible integer estimators, otherwise the float solution is maintained. It follows that indeed the following three cases

can be distinguished:

$\hat{a} \in \Omega_a$	success: correct integer estimation
$\hat{a} \in \Omega \setminus \Omega_a$	failure: incorrect integer estimation
$\hat{a} \notin \Omega$	undecided: ambiguity not fixed to an integer

The corresponding probabilities of success (s), failure (f) and undecidedness (u) are given by:

$$\begin{aligned} P_s &= P(\bar{a} = a) = \int_{\Omega_a} f_{\hat{a}}(x)dx \\ P_f &= \int_{\Omega \setminus \Omega_a} f_{\hat{a}}(x)dx = \int_{\Omega_0} f_{\check{\epsilon}}(x)dx - \int_{\Omega_a} f_{\hat{a}}(x)dx \\ P_u &= 1 - P_s - P_f = 1 - \int_{\Omega_0} f_{\check{\epsilon}}(x)dx \end{aligned} \tag{5.4}$$

These probabilities are referred to as success rate, fail rate, and undecided rate respectively.

5.1.2 Distribution functions

From the definition of the new estimator in (5.2) it follows that the distribution function will take a hybrid form, which means that it is a combination of a probability mass function for all values within Ω, and a probability density function otherwise:

$$\begin{aligned} f_{\bar{a}}(x) &= \begin{cases} f_{\hat{a}}(x) & \text{if } x \in \mathbb{R}^n \setminus \{\Omega\} \\ \int_{\Omega_z} f_{\hat{a}}(y)dy = P(\bar{a} = z) & \text{if } x = z,\ z \in \mathbb{Z}^n \\ 0 & \text{otherwise} \end{cases} \\ &= f_{\hat{a}}(x)\bar{\omega}(x) + \sum_{z \in \mathbb{Z}^n} \int_{\Omega_z} f_{\hat{a}}(y)dy\delta(x - z) \end{aligned} \tag{5.5}$$

where $\delta(x)$ is the impulse function, $\bar{\omega}(x) = (1 - \sum_{z \in \mathbb{Z}^n} \omega_z(x))$ is the indicator function of $\mathbb{R}^n \setminus \Omega$. Note that $P(\bar{a} = z) \leq P(\check{a} = z)$.

Also the distribution function of the corresponding baseline estimator $\bar{b}$ is different compared to the PDF of $\check{b}$:

$$f_{\bar{b}}(x) = \sum_{z \in \mathbb{Z}^n} f_{\hat{b}|\hat{a}}(x|z)P(\bar{a} = z) + \int_{\mathbb{R}^n \setminus \{\Omega\}} f_{\hat{b}|\hat{a}}(x|y)f_{\hat{a}}(y)dy \tag{5.6}$$

In the 1-D case the aperture space is built up of intervals around the integers. As an example, it is assumed that $a = 0$ and $\Omega_a = [-0.2, 0.2]$ and thus $\Omega_z = [-0.2+z, 0.2+z]$. The top panel of figure 5.1 shows the PDF of $\hat{a}$ in black, the PMF of $\check{a}$ in dark grey,

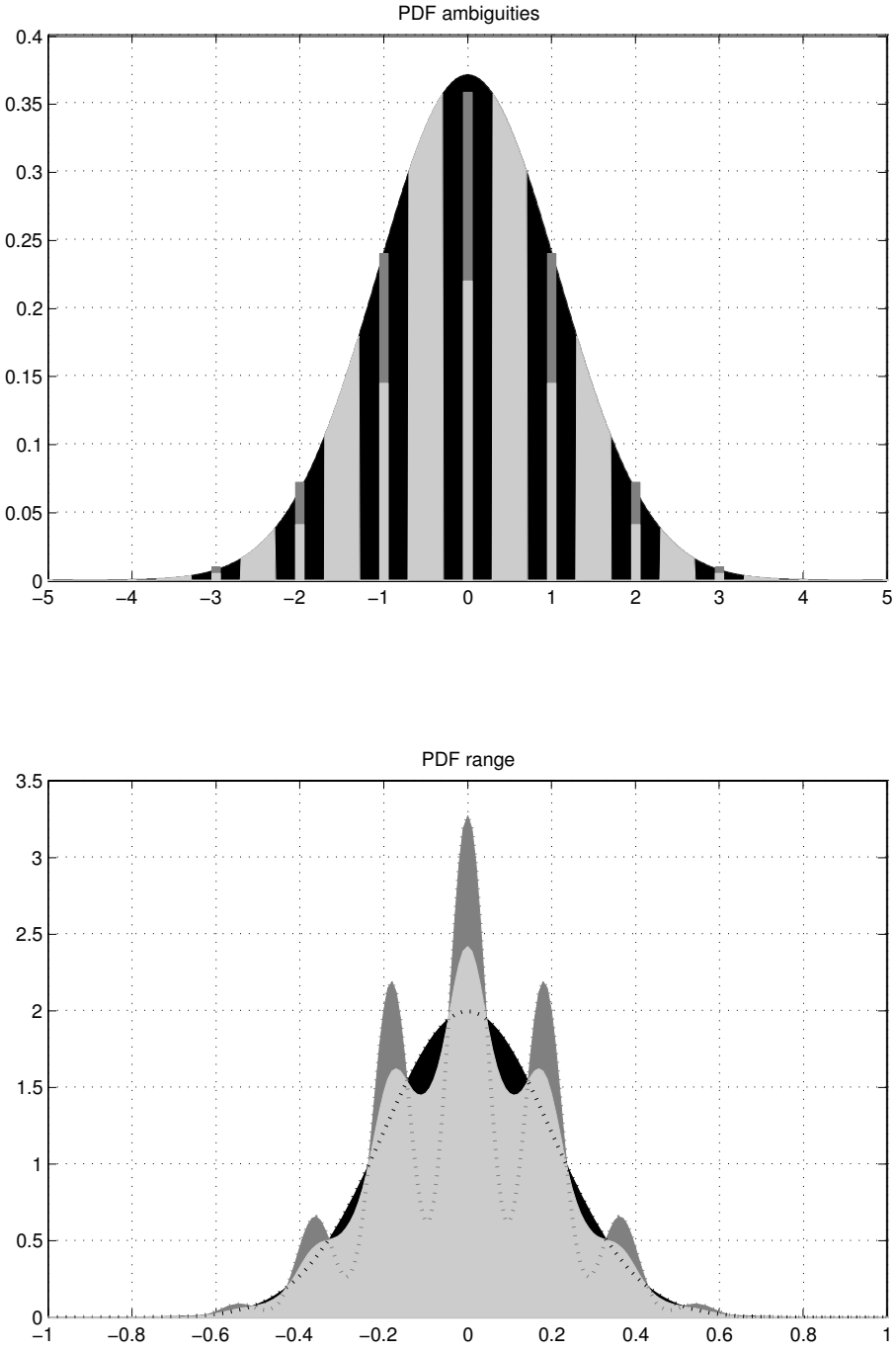

Figure 5.1: 1-D Example of the distribution functions of the different ambiguity estimators (top) and baseline estimators (bottom). Black: float; Dark grey: integer rounding; Light grey: IA.

and the hybrid distribution of $\bar{a}$ in light grey. The bars indicate probability masses. The bottom panel in the figure shows the corresponding distributions of the three baseline estimators. It can be seen that the PDF of $\bar{a}$ always falls in between the PDFs of $\hat{b}$ and $\check{b}$.

5.1.3 IA estimation

It is now possible to design different IA estimators by defining the size and shape of the aperture pull-in regions Ω_z. In classical hypothesis testing theory the size of an acceptance region is determined by the choice of the testing parameters: the false alarm rate and the detection power. However, in the case of integer ambiguity resolution it is not obvious how to choose these parameters. It is especially important that the probability of incorrect ambiguity fixing is small. Therefore the concept of *Integer Aperture estimation with a fixed fail rate* is introduced. This means that the *size* of the aperture space is determined by the condition that the fail rate is equal to or lower than a fixed value. At the same time the *shape* of the aperture pull-in regions should preferably be chosen such that the success rate is still as high as possible. In the following sections several examples of IA estimators are presented.

Note that if all probability mass of $\hat{a}$ is located within the pull-in region S_a the fail rate is automatically equal to zero, and thus the IA solution will always be equal to the fixed solution.

It is very important to note that Integer Aperture estimation with a fixed fail rate is an *overall* approach of integer estimation and validation, and allows for an exact and overall probabilistic evaluation of the solution. With the traditional approaches, e.g. the ratio test applied with a fixed critical value, this is not possible. Two important probabilistic measures are the fail rate, which will never exceed the user-defined threshold, and the probability that $\check{a} = a$ if $\bar{a} = z$:

$$\begin{aligned} P_{s|\bar{a}=z} &= P(\check{a} = a|\bar{a} = z) \\ &= \frac{P(\check{a} = a, \bar{a} = z)}{P(\bar{a} = z)} \\ &= \frac{P_s}{P_s + P_f} \end{aligned} \tag{5.7}$$

The success and fail rate, P_s and P_f, from equation (5.4) are used. Note that if $P_f << P_s$:

$$P_{s|\bar{a}=z} \approx 1 \tag{5.8}$$

This is an important result, since it means that with IA estimation it is not only guaranteed that the fail rate will be below a given threshold, but also that if IA estimation results in a fix, the probability that the integer solution is correct will be close to one if $P_f << P_s$. Only if P_s is small, this is not necessarily the case, but then the probability that the solution will be fixed is also very small, since this probability equals $P_s + P_f$.

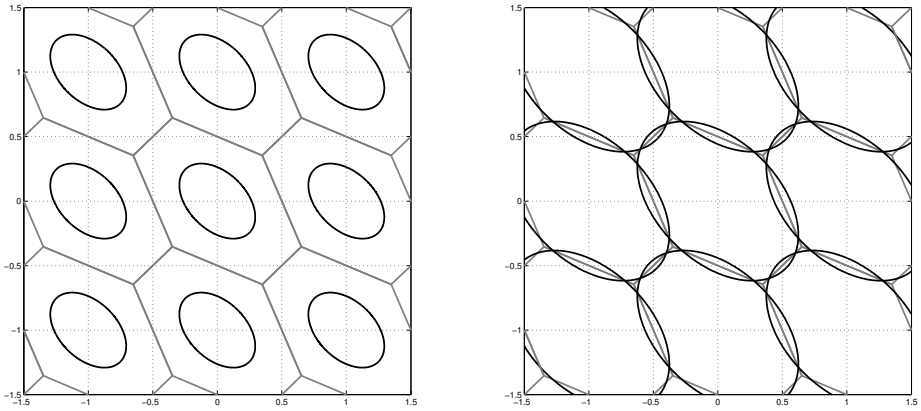

Figure 5.2: Examples of ellipsoidal aperture regions without overlap (left) and with overlap (right).

5.2 Ellipsoidal integer aperture estimation

An easy way to define the aperture space would be to choose the aperture pull-in regions Ω_z as ellipsoidal regions, cf. (Teunissen 2003a).

Let the ellipsoidal region C_0^μ be given as:

$$C_0^\mu = \{x \in \mathbb{R}^n \mid \| x \|^2_{Q_{\hat{a}}} \leq \mu^2\} \tag{5.9}$$

The aperture pull-in regions of the ellipsoidal integer aperture (EIA) estimator are then defined as:

$$\begin{aligned} &\Omega_{0,\mathrm{E}} = S_0 \cap C_0^\mu \\ &\Omega_{z,\mathrm{E}} = \Omega_{0,\mathrm{E}} + z = \{x \in S_z \mid \| x - z \|^2_{Q_{\hat{a}}} \leq \mu^2\},\ \forall z \in \mathbb{Z}^n \end{aligned} \tag{5.10}$$

This shows that the procedure for computing the EIA estimator is rather straightforward. Using the float solution $\hat{a}$, its vc-matrix $Q_{\hat{a}}$, and the aperture parameter μ as input, one only needs to compute the integer least-squares solution $\check{a}$ and verify whether or not the inequality

$$\| \hat{a} - \check{a} \|^2_{Q_{\hat{a}}} \leq \mu^2 \tag{5.11}$$

is satisfied. If the inequality is satisfied then $\bar{a} = \check{a}$, otherwise $\bar{a} = \hat{a}$.

The simple choice of the ellipsoidal criterion (5.11) is motivated by the fact that the squared-norm of a normally distributed random vector is known to have a χ^2-distribution. That is, if $\hat{a}$ is normally distributed then $P(\hat{a} \in C_z^\mu) = P\left(\chi^2\left(n, \lambda_z\right) \leq \mu^2\right)$, with non-centrality parameter $\lambda_z = (z-a)^T Q_{\hat{a}}^{-1}(z-a)$. This implies that one can give *exact* solutions to the fail rate and to the success rate of the EIA estimator, provided the ellipsoidal regions C_z^μ do not overlap. The EIA probabilities of failure, success and

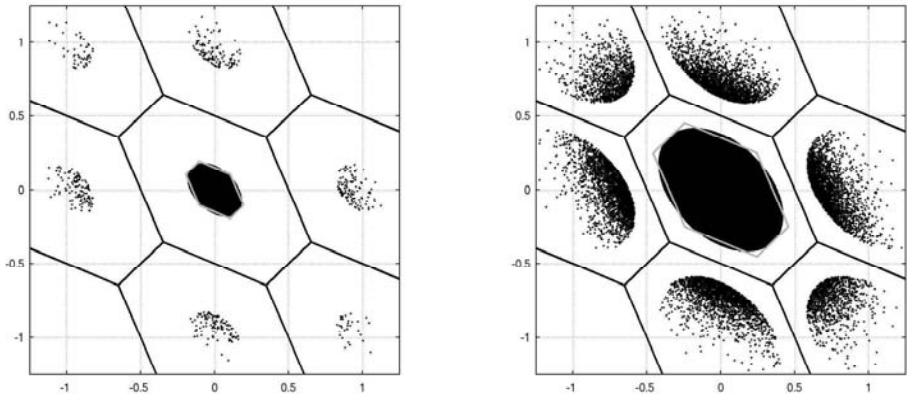

Figure 5.3: 2-D example for EIA estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

undecided are given as:

$$\begin{cases} P_f = \sum\limits_{z \in Z^n \setminus \{0\}} P\left(\chi^2(n, \lambda_z) \leq \mu^2\right) \\ P_s = P\left(\chi^2(n, 0) \leq \mu^2\right) \\ P_u = 1 - P_f - P_s \end{cases} \tag{5.12}$$

The non-centrality parameter is given by $\lambda_z = z^T Q_{\hat{a}}^{-1} z$.

Note that the region for which the EIA estimator outputs an integer as solution is given as $\Omega_{\mathsf{E}} = \cup_{z \in \mathbb{Z}^n} \Omega_{z,\mathsf{E}}$. Since this region is independent of S_z it follows that $\Omega_{\mathsf{E}} = \cup_{z \in \mathbb{Z}^n} C_z^{\mu}$. Thus Ω_{E} equals the union of all integer translated copies of the origin-centered ellipsoidal region C_0^{μ}. Depending on the value chosen for the aperture parameter μ, these ellipsoidal regions may or may not overlap, as illustrated in figure 5.2. The overlap will occur when μ is larger than half times the smallest distance between two integer vectors. Thus when μ is larger than half times the length of the smallest nonzero integer vector, $\mu > \frac{1}{2} \min_{z \in \mathbb{Z}^n \setminus \{0\}} \parallel z \parallel_{Q_{\hat{a}}}$. In that case the probabilities given in equation (5.12) for P_f and P_s become upper bounds. In order to obtain the corresponding lower bounds, one has to replace μ by $\mu' = \frac{1}{2} \min_{z \in \mathbb{Z}^n \setminus \{0\}} \parallel z \parallel_{Q_{\hat{a}}}$, since $C_z^{\mu'} \subset \Omega_{z,\mathsf{E}}$ and $C_z^{\mu'} \subset S_z$.

In the absence of overlap exact evaluation of the EIA estimator is possible. The steps for computing and evaluating the EIA estimator are as follows. Depending on the application the user will usually start by requiring a certain small and fixed value for the fail rate P_f. Hence, the user will start by choosing the aperture parameter μ such that it meets his requirement on P_f. For a certain value for P_f, the value of the aperture parameter μ needs to be chosen smaller when the precision of the ambiguities becomes poorer. This can be understood by looking at the non-centrality parameter $\lambda_z = z^T Q_{\hat{a}}^{-1} z$, since a poorer precision will result in a smaller value for λ_z and thus in a larger value for the probability $P\left(\chi^2(n, \lambda_z) \leq \mu^2\right)$.

2-D example

Figure 5.3 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

5.3 Ratio test, difference test and projector test

In section 3.5 several integer validation tests as proposed in the past were described. It was explained that the test criteria cannot be derived using classical hypothesis tests. However, it can be shown that the ratio test (3.78), the difference test (3.86), and the projector test, on which the test statistics (3.87) and (3.89) are based, are all examples of Integer Aperture estimators, cf. (Teunissen 2003e; Verhagen and Teunissen 2004a).

5.3.1 Ratio test is an IA estimator

The ratio test of (3.78) is a very popular validation test in practice. Instead of (3.78) the 'inverse' of the ratio test is used here, i.e. accept $\check{a}$ if:

$$\frac{\|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}}}{\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}} \leq \mu$$

This guarantees that the critical value is bounded as $0 < \mu \leq 1$, since by definition $\|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}} \leq \|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}$.

It can now be shown that integer estimation in combination with the ratio test is an IA estimator. The acceptance region is given as:

$$\Omega_{\mathsf{R}} = \{x \in \mathbb{R}^n \mid \|x - \check{x}\|^2_{Q_{\hat{a}}} \leq \mu \|x - \check{x}_2\|^2_{Q_{\hat{a}}},\ 0 < \mu \leq 1\} \tag{5.13}$$

with $\check{x}$ and $\check{x}_2$ the best and second-best ILS estimator of x. Let $\Omega_{z,\mathsf{R}} = \Omega_{\mathsf{R}} \cap S_z$, i.e. $\Omega_{z,\mathsf{R}}$ is the intersection of Ω_{R} with the ILS pull-in region as defined in (3.16). Then all conditions of (5.1) are fulfilled, since:

$$\begin{cases} \Omega_{0,\mathsf{R}} & = \{x \in \mathbb{R}^n \mid \|x\|^2_{Q_{\hat{a}}} \leq \mu \|x - z\|^2_{Q_{\hat{a}}},\ \forall z \in \mathbb{Z}^n \setminus \{0\}\} \\ \Omega_{z,\mathsf{R}} & = \Omega_{0,\mathsf{R}} + z,\ \forall z \in \mathbb{Z}^n \\ \Omega_{\mathsf{R}} & = \bigcup_{z \in \mathbb{Z}^n} \Omega_{z,\mathsf{R}} \end{cases} \tag{5.14}$$

The proof was given in Teunissen (2003e) and Verhagen and Teunissen (2004a) and is

as follows:

$$\begin{aligned}
\Omega_{z,\mathsf{R}} &= \Omega_{\mathsf{R}} \cap S_z \\
&= \{x \in S_z \mid \|x-\check{x}\|^2_{Q_{\hat{a}}} \leq \mu \|x-\check{x}_2\|^2_{Q_{\hat{a}}},\, 0<\mu\leq 1\} \\
&= \{x \in S_z \mid \|x-\check{x}\|^2_{Q_{\hat{a}}} \leq \mu \|x-u\|^2_{Q_{\hat{a}}},\, \forall u \in \mathbb{Z}^n \setminus \{\check{x}\}, 0<\mu\leq 1\} \\
&= \{x \in S_z \mid \|x-z\|^2_{Q_{\hat{a}}} \leq \mu \|x-u\|^2_{Q_{\hat{a}}},\, \forall u \in \mathbb{Z}^n \setminus \{z\}, 0<\mu\leq 1\} \\
&= \{x \in \mathbb{R}^n \mid \|x-z\|^2_{Q_{\hat{a}}} \leq \mu \|x-u\|^2_{Q_{\hat{a}}},\, \forall u \in \mathbb{Z}^n \setminus \{z\}, 0<\mu\leq 1\} \\
&= \{x \in \mathbb{R}^n \mid \|y\|^2_{Q_{\hat{a}}} \leq \mu \|y-v\|^2_{Q_{\hat{a}}},\, \forall v \in \mathbb{Z}^n \setminus \{0\},\ x=y+z, 0<\mu\leq 1\} \\
&= \Omega_{\mathsf{R}} \cap S_0 + z \\
&= \Omega_{0,\mathsf{R}} + z
\end{aligned}$$

The first two equalities follow from the definition of the ratio test, and the third from $\|\hat{x}-\check{x}\|^2_{Q_{\hat{a}}} \leq \|\hat{x}-\check{x}_2\|^2_{Q_{\hat{x}}} \leq \|\hat{x}-u\|^2_{Q_{\hat{a}}}, \forall u \in \mathbb{Z}^n \setminus \{\check{x}\}$. The fourth equality follows since $\check{x}=z$ is equivalent to $x \in S_z$, and the fifth equality from the fact that $0<\mu\leq 1$. The last equalities follow from a change of variables and the definition of $\Omega_{0,\mathsf{R}} = \Omega_{\mathsf{R}} \cap S_0$. Finally note that

$$\bigcup_{z\in\mathbb{Z}^n} \Omega_{z,\mathsf{R}} = \bigcup_{z\in\mathbb{Z}^n} (\Omega_{\mathsf{R}} \cap S_z) = \Omega_{\mathsf{R}} \cap (\bigcup_{z\in\mathbb{Z}^n} S_z) = \Omega_{\mathsf{R}}$$

This ends the proof of (5.14).

It has thus been shown that the acceptance region of the ratio test consists of an infinite number of regions, each one is an integer translated copy of $\Omega_{0,\mathsf{R}} \subset S_0$. The acceptance region plays the role of the aperture space, and μ plays the role of an aperture parameter. The ratio test will be referred to as the RTIA estimator.

It can now be shown how the aperture pull-in regions for the ratio test are 'constructed', cf. (Verhagen and Teunissen 2004a). From the definition of Ω_0 the following can be derived:

$$\begin{aligned}
\Omega_{0,\mathsf{R}} : \|x\|^2_{Q_{\hat{a}}} &\leq \mu \|x-z\|^2_{Q_{\hat{a}}}, \quad \forall z \in \mathbb{Z}^n \setminus \{0\},\ 0<\mu\leq 1 \\
&\iff \|\frac{1-\mu}{\mu}x + z\|^2_{Q_{\hat{a}}} \leq \frac{1}{\mu}\|z\|^2_{Q_{\hat{a}}}, \quad \forall z \in \mathbb{Z}^n \setminus \{0\} \\
&\iff \|x + \frac{\mu}{1-\mu}z\|^2_{Q_{\hat{a}}} \leq \frac{\mu}{(1-\mu)^2}\|z\|^2_{Q_{\hat{a}}}, \quad \forall z \in \mathbb{Z}^n \setminus \{0\}
\end{aligned} \tag{5.15}$$

This shows that the aperture pull-in region is equal to the intersection of all ellipsoids with centers $-\frac{\mu}{1-\mu}z$ and size governed by $\frac{\sqrt{\mu}}{1-\mu}\|z\|_{Q_{\hat{a}}}$. In fact, the intersection region is only determined by the adjacent integers, since $\forall u \in Z^n$ that are not adjacent, (5.15) is always fulfilled $\forall x \in S_0$. It follows namely from equations (5.13) and (5.14) that for all $x \in S_0$ on the boundary of Ω_0:

$$\|x\|^2_{Q_{\hat{a}}} = \mu\|x-c\|^2_{Q_{\hat{a}}},\ c = \arg\min_{z\in\mathbb{Z}^n\setminus\{0\}} \|x-z\|^2_{Q_{\hat{a}}}$$

So, c is the second-closest integer and must be adjacent. Therefore the following is always true:

$$\|x\|^2_{Q_{\hat{a}}} = \mu\|x-c\|^2_{Q_{\hat{a}}} \leq \mu\|x-z\|^2_{Q_{\hat{a}}},\ \forall z \in \mathbb{Z}^n \setminus \{0,c\}$$

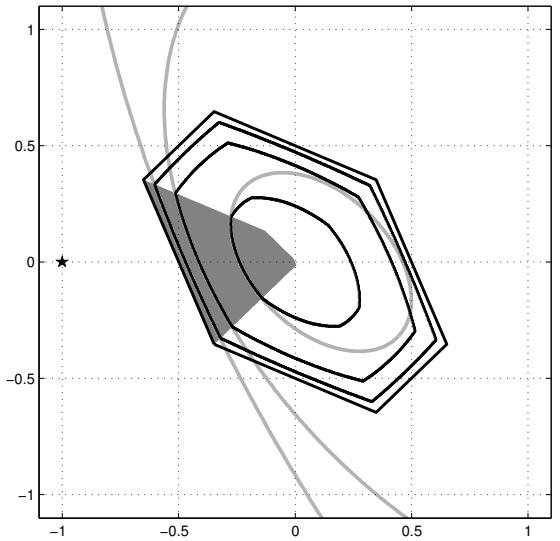

Figure 5.4: Construction of RTIA aperture pull-in region. For all x that fall in the grey region, the second-closest integer $z_2 = [-1\ 0]^T$. The boundary of the aperture pull-in region within this grey region is then equal to the ellipsoid with center $-\frac{\mu}{1-\mu}z_2$ and size governed by $\frac{\sqrt{\mu}}{1-\mu}\|z_2\|_{Q_{\hat{a}}}$. Examples are shown for μ is 0.3, 0.5, and 0.8, respectively. Ellipsoids are grey; aperture pull-in regions black.

so that the integers z do not have any influence on the boundary of the aperture pull-in region.

This means that the integer least-squares pull-in region S_0 can be split up into sectors, all having another integer c as second-closest, i.e. $\|x\|^2_{Q_{\hat{a}}} \leq \|x-c\|^2_{Q_{\hat{x}}} \leq \|x-z\|^2_{Q_{\hat{a}}}, \forall z \in \mathbb{Z}^n \setminus \{0\}$. Within a sector, the aperture pull-in region is then equal to the intersection of the sector with the ellipsoid with center $-\frac{\mu}{1-\mu}c$ and size governed by $\frac{\sqrt{\mu}}{1-\mu}\|c\|_{Q_{\hat{a}}}$. This is illustrated for the 2-D case in figure 5.4.

It can be seen that especially for larger μ, the shape of the aperture pull-in region starts to resemble that of the ILS pull-in region. The reason is that for larger μ the size of the ellipsoid increases, and the center of the ellipsoid is further away, but in the direction of $[0, -c]$, and the ellipsoid has the same orientation as the ILS pull-in region.

2-D example

Figure 5.5 shows two 2-D examples of the construction of the aperture pull-in regions. For a diagonal matrix the ILS pull-in region becomes square and there are only four adjacent integers. It can be seen in the left panel that indeed only these four integers determine the shape of the aperture pull-in region. For the vc-matrix $Q_{\hat{z}\_02\_01}$ there are six adjacent integers that determine the shape.

Figure 5.6 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

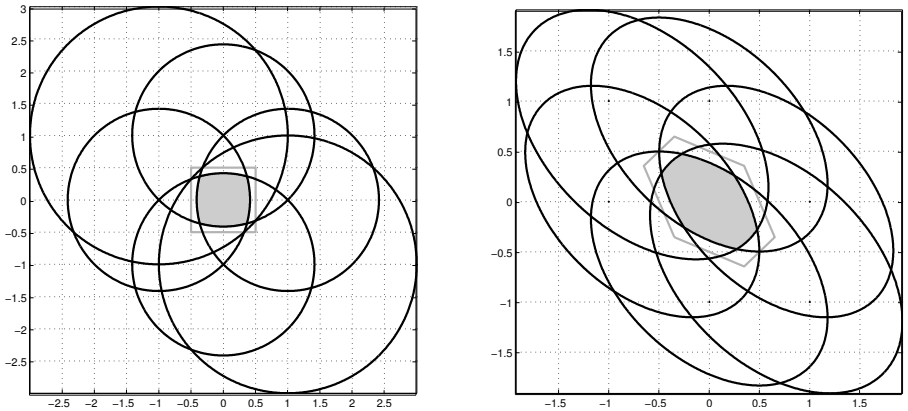

Figure 5.5: Construction of RTIA aperture pull-in region (grey); examples for different covariance matrices. Left: identity matrix, $\mu = 0.5$; Right: $Q_{\hat{z}\_02\_01}$, $\mu = 0.4$.

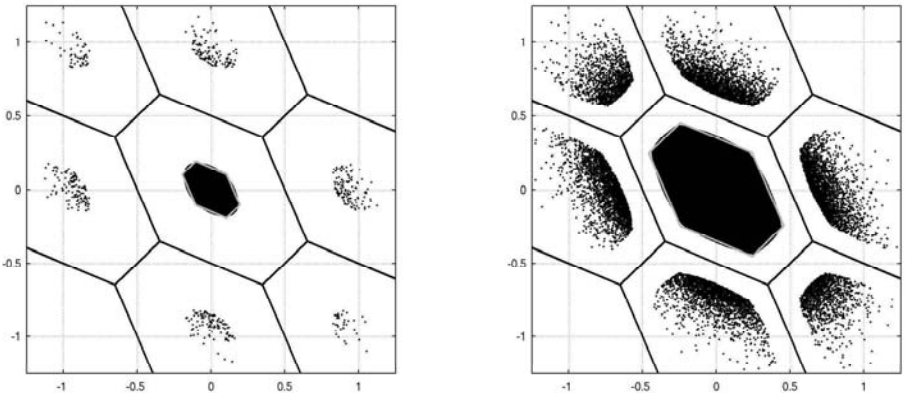

Figure 5.6: 2-D example for RTIA estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

5.3.2 F-Ratio test

With the ratio test considered above the test is always accepted or rejected for a certain value of the ambiguity residuals and for a certain μ. With the F-ratio test (3.77), however, this is not the case. The test statistic namely depends on the residuals of the complete float solution and on the ambiguity residuals. The fixed solution is accepted if:

$$\frac{\|\hat{e}\|^2_{Q_y} + \|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}}}{\|\hat{e}\|^2_{Q_y} + \|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}} \leq \mu \quad \text{with } 0 < \mu \leq 1$$

Still, it can be shown that the procedure underlying this test is a member from the class of IA estimators. An integer perturbation of the observations $y + Az$, with $z \in \mathbb{Z}^n$ does namely not propagate into the least-squares residuals, and $\hat{a}$, $\check{a}$, and $\check{a}_2$ all undergo the

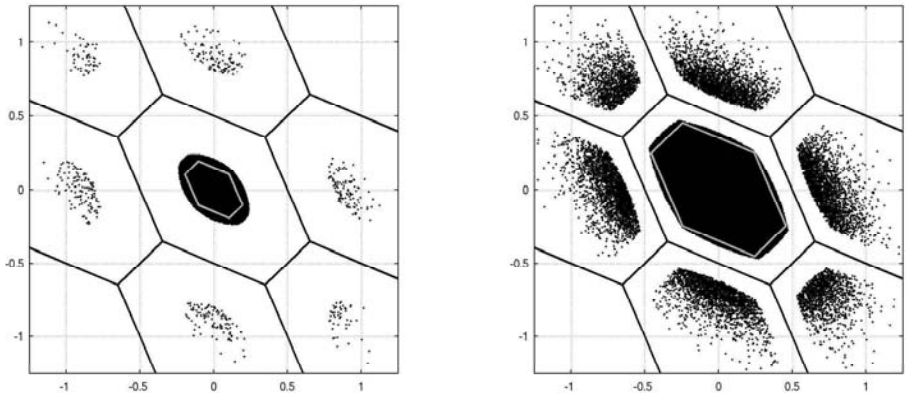

Figure 5.7: 2-D example for F-RTIA estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

same change, so that both $\hat{a} - \check{a}$ and $\hat{a} - \check{a}_2$ are invariant under the integer perturbation. This IA estimator will be referred to as the F-RTIA estimator.

Similarly as for the RTIA estimator it can be shown how the aperture pull-in regions of the F-RTIA estimator are constructed, (Verhagen and Teunissen 2004a):

$$\begin{aligned}\Omega_{0,\mathrm{R}} : \|x\|^2_{Q_{\hat{a}}} + \|\hat{e}\|^2_{Q_y} &\leq \mu(\|x-z\|^2_{Q_{\hat{a}}} + \|\hat{e}\|^2_{Q_y}), \quad \forall z \in \mathbb{Z}^n \setminus \{0\},\ 0 < \mu \leq 1 \\ &\iff \|x\|^2_{Q_{\hat{a}}} \leq \mu\|x-z\|^2_{Q_{\hat{a}}} + (\mu-1)\|\hat{e}\|^2_{Q_y} \\ &\iff \|x + \frac{\mu}{1-\mu}z\|^2_{Q_{\hat{a}}} \leq \frac{\mu}{(1-\mu)^2}\|z\|^2_{Q_{\hat{a}}} - \|\hat{e}\|^2_{Q_y}, \quad \forall z \in \mathbb{Z}^n \setminus \{0\} \end{aligned} \tag{5.16}$$

So, Ω_0 is equal to the intersection of all ellipsoids with centers $-\frac{\mu}{1-\mu}z$ and size governed by the square root of the term on the right-hand side of the last inequality in equation (5.16). The shape of the aperture pull-in region is identical to that of the RTIA estimator. Obviously, the difference with the RTIA estimator is that the sizes depend on the sample value of $\hat{e}$. For the same value of μ they will be smaller than or the same size as the ellipsoids that determine the RTIA aperture pull-in region.

2-D example

Figure 5.7 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates. Since the size of the aperture pull-in regions depends on the sample values of $\hat{e}$, the point clouds give an indication of the *maximum* size of the aperture pull-in regions; in general they will be smaller for a certain sample.

5.3.3 Difference test is an IA estimator

The difference test (3.86) as proposed by (Tiberius and De Jonge 1995) leads to acceptance of $\check{a}$ if:

$$\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}} - \|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}} \geq \mu \tag{5.17}$$

The acceptance region of this test is given as:

$$\Omega_{\mathsf{D}} = \{x \in \mathbb{R}^n \mid \|x - \check{x}\|^2_{Q_{\hat{a}}} \leq \|x - \check{x}_2\|^2_{Q_{\hat{a}}} - \mu\} \tag{5.18}$$

Let $\Omega_{z,\mathsf{D}} = \Omega_{\mathsf{D}} \cap S_z$. Then:

$$\begin{cases} \Omega_{0,\mathsf{D}} & = \{x \in \mathbb{R}^n \mid \|x\|^2_{Q_{\hat{a}}} \leq \|x - z\|^2_{Q_{\hat{a}}} - \mu,\ \forall z \in \mathbb{Z}^n \setminus \{0\}\} \\ \Omega_{z,\mathsf{D}} & = \Omega_{0,\mathsf{D}} + z,\ \forall z \in \mathbb{Z}^n \\ \Omega_{\mathsf{D}} & = \bigcup_{z \in \mathbb{Z}^n} \Omega_{z,\mathsf{D}} \end{cases} \tag{5.19}$$

The proof given in Teunissen (2003e) and Verhagen and Teunissen (2004a) is as follows:

$$\begin{aligned} \Omega_{z,\mathsf{D}} &= \Omega_{\mathsf{D}} \cap S_z \\ &= \{x \in S_z \mid \|x - \check{x}\|^2_{Q_{\hat{a}}} \leq \|x - \check{x}_2\|^2_{Q_{\hat{a}}} - \mu\} \\ &= \{x \in S_z \mid \|x - \check{x}\|^2_{Q_{\hat{a}}} \leq \|x - u\|^2_{Q_{\hat{a}}} - \mu,\ \forall u \in \mathbb{Z}^n \setminus \{\check{x}\}\} \\ &= \{x \in S_z \mid \|x - z\|^2_{Q_{\hat{a}}} \leq \|x - u\|^2_{Q_{\hat{a}}} - \mu,\ \forall u \in \mathbb{Z}^n \setminus \{z\}\} \\ &= \{x \in \mathbb{R}^n \mid \|x - z\|^2_{Q_{\hat{a}}} \leq \|x - u\|^2_{Q_{\hat{a}}} - \mu,\ \forall u \in \mathbb{Z}^n \setminus \{z\}\} \\ &= \{x \in \mathbb{R}^n \mid \|y\|^2_{Q_{\hat{a}}} \leq \|y - v\|^2_{Q_{\hat{a}}} - \mu,\ \forall v \in \mathbb{Z}^n \setminus \{0\}, x = y + z\} \\ &= \Omega_{\mathsf{D}} \cap S_0 + z \\ &= \Omega_{0,\mathsf{D}} + z \end{aligned}$$

The first two equalities follow from the definition of the difference test, and the third from $\|\hat{x} - \check{x}\|^2_{Q_{\hat{x}}} \leq \|\hat{x} - \check{x}_2\|^2_{Q_{\hat{a}}} \leq \|\hat{x} - u\|^2_{Q_{\hat{a}}}, \forall u \in \mathbb{Z}^n \setminus \{\check{x}\}$. The fourth equality follows since $\check{x} = z$ is equivalent to $x \in S_z$, and the fifth equality from the fact that $\mu \geq 0$. The last equalities follow from a change of variables and the definition of $\Omega_{0,\mathsf{D}} = \Omega_{\mathsf{D}} \cap S_0$. Finally note that

$$\bigcup_{z \in \mathbb{Z}^n} \Omega_{z,\mathsf{D}} = \bigcup_{z \in \mathbb{Z}^n} (\Omega_{\mathsf{D}} \cap S_z) = \Omega_{\mathsf{D}} \cap (\bigcup_{z \in \mathbb{Z}^n} S_z) = \Omega_{\mathsf{D}}$$

This ends the proof of (5.19).

It has been shown that also in the case of the difference test the acceptance region consists of an infinite number of integer translated copies of a subset of the ILS pull-in region S_0. And thus, the difference test is an IA estimator with aperture parameter μ. It will be referred to as the DTIA estimator.

In a similar way as for the ratio test aperture pull-in region, it is possible to show how the aperture pull-in region of the difference test is constructed, see (Verhagen and Teunissen

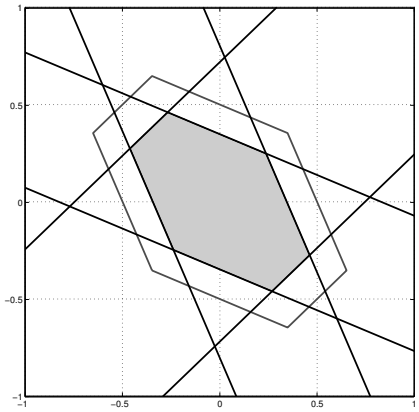

Figure 5.8: Construction of DTIA aperture pull-in region (grey) for $Q_{\hat{z}\_02\_01}$.

2004a). From the definition of $\Omega_{0,\mathrm{D}}$ follows that:

$$\begin{aligned}\Omega_{0,\mathrm{D}} : \|x\|^2_{Q_{\hat{a}}} &\leq \|x-z\|^2_{Q_{\hat{a}}} - \mu, \quad \forall z \in \mathbb{Z}^n \setminus \{0\},\ \mu \geq 0 \\ &\iff x^T Q^{-1}_{\hat{a}} z \leq \frac{1}{2}(\|z\|^2_{Q_{\hat{a}}} - \mu) \\ &\iff \frac{z^T Q^{-1}_{\hat{a}} x}{\|z\|_{Q_{\hat{a}}}} \leq \frac{\|z\|^2_{Q_{\hat{a}}} - \mu}{2\|z\|_{Q_{\hat{a}}}}, \quad \forall z \in \mathbb{Z}^n \setminus \{0\} \end{aligned} \tag{5.20}$$

On the left-hand side of equation (5.20) the orthogonal projection of x onto the direction z can be recognized. This shows that the aperture pull-in region $\Omega_{0,\mathrm{D}}$ is constructed as intersecting half-spaces which are bounded by the planes orthogonal to z and passing through the points $\frac{1}{2}(1 - \frac{\mu}{\|z\|^2_{Q_{\hat{a}}}})z$.

The construction of $\Omega_{0,\mathrm{D}}$ is thus very similar to that of the ILS pull-in region S_0, see equation (3.17), which is also constructed of half-spaces bounded by the planes orthogonal to z. The difference is that these planes pass through the midpoint $\frac{1}{2}z$, whereas in the case of the difference test this point depends on the distance $\|z\|^2_{Q_{\hat{a}}}$ and on μ. This implies that the difference test aperture pull-in region is a down-scaled version of the ILS pull-in region, but that the scaling is different in the various directions.

2-D example

Figure 5.8 shows a 2-D example of the construction of the aperture pull-in region. The black lines are the planes orthogonal to c and passing through the point $\frac{1}{2}(1 - \frac{\mu}{\|c\|^2_{Q_{\hat{a}}}})c$, where the c are the six adjacent integers.

Figure 5.9 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

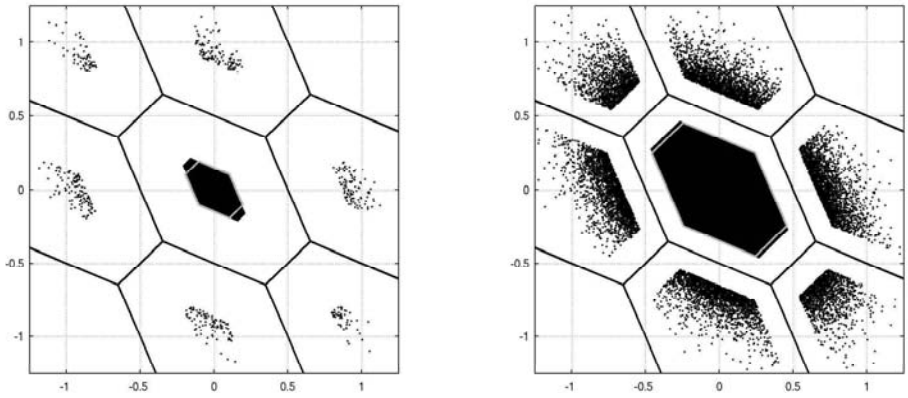

Figure 5.9: 2-D example for DTIA estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

5.3.4 Projector test is an IA estimator

The tests defined in Wang et al. (1998) and in Han (1997) are both based on testing:

$$H_0 : E\{\hat{a}\} = \check{a} \text{ versus } H_a : E\{\hat{a}\} = \check{a} + c\nabla, \text{ with } c = \frac{\check{a}_2 - \check{a}}{\|\check{a}_2 - \check{a}\|_{Q_{\hat{a}}}}$$

$\check{a}_2$ is the second-best integer candidate. H_a can be tested against H_0 using the w-test of equation (2.82). The fixed solution is accepted if:

$$\left| \frac{(\check{a}_2 - \check{a})^T Q_{\hat{a}}^{-1} (\hat{a} - \check{a})}{\|\check{a}_2 - \check{a}\|_{Q_{\hat{a}}}} \right| \leq \mu \tag{5.21}$$

This test is referred to as the projector test since the term on the left-hand side of equation (5.21) equals a projector which projects $\hat{a} - \check{a}$ orthogonally on the direction of $\check{a}_2 - \check{a}$, in the metric of $Q_{\hat{a}}$. Furthermore note that from equation (3.90) it follows that:

$$\left| \frac{(\check{a}_2 - \check{a})^T Q_{\hat{a}}^{-1} (\hat{a} - \check{a})}{\|\check{a}_2 - \check{a}\|_{Q_{\hat{a}}}} \right| \leq \frac{1}{2} \|\check{a}_2 - \check{a}\|_{Q_{\hat{a}}} \tag{5.22}$$

This implies that the fixed solution is always accepted if $\mu \geq \frac{1}{2}\|\check{a}_2 - \check{a}\|_{Q_{\hat{a}}}$.

The acceptance region of the projector test is given as:

$$\Omega_{\mathsf{P}} = \{x \in \mathbb{R}^n \mid (\check{x} - \check{x}_2)^T Q_{\hat{a}}^{-1} (x - \check{x}) \leq \mu \|\check{x} - \check{x}_2\|_{Q_{\hat{a}}}\} \tag{5.23}$$

In a similar way as for the difference test it can be proven that

$$\begin{cases} \Omega_{0,\mathsf{P}} & = \{x \in S_0 \mid c^T Q_{\hat{a}}^{-1} x \leq \mu \|c\|_{Q_{\hat{a}}},\, c = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|^2_{Q_{\hat{a}}}\} \\ \Omega_{z,\mathsf{P}} & = \Omega_{0,\mathsf{P}} + z,\ \forall z \in \mathbb{Z}^n \\ \Omega_{\mathsf{P}} & = \bigcup_{z \in \mathbb{Z}^n} \Omega_{z,\mathsf{P}} \end{cases} \tag{5.24}$$

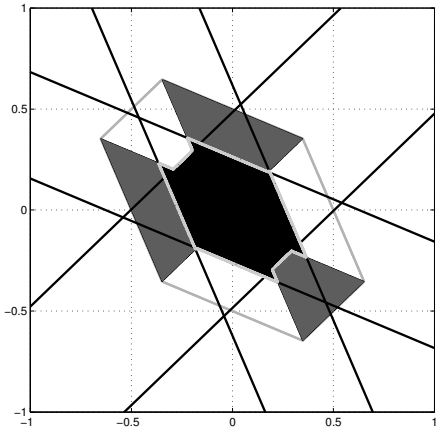

Figure 5.10: Construction of projector test aperture pull-in region for $Q_{\hat{z},02\_01}$, $\mu = 1$.

And thus also the projector test is an IA estimator, and will be referred to as the PTIA estimator.

From equation (5.24) follows that $\Omega_{0,\mathrm{P}}$ is bounded by the planes orthogonal to c and passing through μc, and these planes themselves are bounded by the condition that $c = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|^2_{Q_{\hat{a}}}$ for all x on the plane, see (Verhagen and Teunissen 2004a).

2-D example

Figure 5.10 shows a 2-D example of the construction of the aperture pull-in region Ω_0. The black lines are the planes orthogonal to c and passing through μc. For the construction of $\Omega_{0,\mathrm{P}}$, these planes are bounded by the condition that $c = \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|x - z\|^2_{Q_{\hat{a}}}$ for all $x \in \Omega_0$. Therefore, also sectors within the ILS pull-in region are shown as alternating grey and white regions, with the sectors containing all x with a certain integer c as second closest integer. The region $\Omega_{0,\mathrm{P}}$ follows as the region bounded by the intersection of the black lines and the sectors, *and* by the boundaries of the sectors. This results in the strange shape of the aperture pull-in regions in the direction of the vertices of the ILS pull-in region.

Figure 5.11 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

5.4 Integer Aperture Bootstrapping and Least-Squares

A straightforward choice of the aperture pull-in regions would be to choose them as down-scaled versions of the pull-in regions S_z corresponding to a certain admissible integer estimator. This approach is applied to the bootstrapping and integer least-squares estimators.

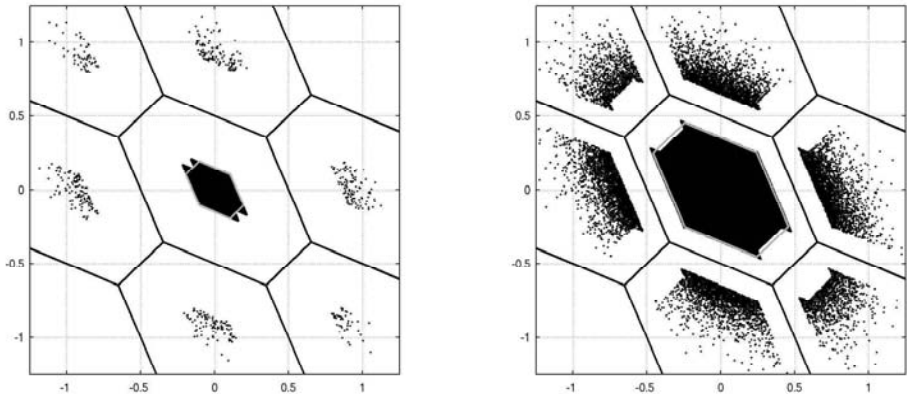

Figure 5.11: 2-D example for PTIA estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

5.4.1 Integer Aperture Bootstrapping

Integer bootstrapping is an easy-to-evaluate integer estimator, for which exact evaluation of the success rate is possible, see section 3.1.2. It will be shown that this is also the case by defining the IA bootstrapped (IAB) estimator by choosing the aperture pull-in regions as down-scaled versions of the bootstrapped pull-in region $S_{z,\mathsf{B}}$ as defined in equation (3.12), cf. (Teunissen 2004a):

$$\Omega_{z,\mathsf{B}} = \mu S_{z,\mathsf{B}},\ \forall z \in \mathbb{Z}^n \tag{5.25}$$

with

$$\begin{cases} \mu S_{z,\mathsf{B}} = \{x \in \mathbb{R}^n \mid \frac{1}{\mu}(x - z) \in S_{0,\mathsf{B}}\} \\ S_{0,\mathsf{B}} = \{x \in \mathbb{R}^n \mid |c_i^T L^{-1} x| \leq \frac{1}{2},\ i = 1, \ldots, n\} \end{cases} \tag{5.26}$$

The aperture parameter is μ, $0 \leq \mu \leq 1$.

Like for the original bootstrapped estimator, computation of the IA bootstrapped solution is quite simple. First the bootstrapped solution for the given float solution $\hat{a}$ is computed. The aperture pull-in region $\Omega_{\check{a}_\mathsf{B},\mathsf{B}}$ is then identified, and one needs to verify whether or not the float solution resides in it. This is equivalent to verifying whether or not

$$\frac{1}{\mu}(\hat{a} - \check{a}_\mathsf{B}) = \frac{1}{\mu}\check{\epsilon}_\mathsf{B} \in S_{0,\mathsf{B}} \tag{5.27}$$

This can be done by applying the bootstrapped procedure again, but now applied to the up-scaled version of the ambiguity residuals, $\frac{1}{\mu}\check{\epsilon}_\mathsf{B}$. If the outcome is the zero vector, then $\bar{a} = \check{a}_\mathsf{B}$, otherwise $\bar{a} = \hat{a}$.

It is now possible to derive closed-form expressions for the computation of the IAB

success rate and fail rate. From equation (5.4) it follows that:

$$\begin{aligned}
P_f &= \sum_{z\in\mathbb{Z}^n\setminus\{a\}} \int_{\mu S_{z,\mathsf{B}}} f_{\hat{a}}(x)dx \\
&= \sum_{z\in\mathbb{Z}^n\setminus\{0\}} \int_{S_{0,\mathsf{B}}} \frac{1}{(2\pi)^{\frac{1}{2}n}\sqrt{|\frac{1}{\mu^2}Q_{\hat{a}}|}} \exp\{-\frac{1}{2}\parallel x+\frac{1}{\mu}z\parallel^2_{\frac{1}{\mu^2}Q_{\hat{a}}}\}dx \\
&= \sum_{z\in\mathbb{Z}^n\setminus\{0\}} \int_{F^{-1}(S_{0,\mathsf{B}})} \frac{1}{(2\pi)^{\frac{1}{2}n}\sqrt{|\frac{1}{\mu^2}D|}} \exp\{-\frac{1}{2}\parallel y+\frac{1}{\mu}L^{-1}z\parallel^2_{\frac{1}{\mu^2}D}\}dy
\end{aligned}$$

where for the third equality the transformation $F : x = Ly$ is used, with L the unit lower triangular matrix of $Q_{\hat{a}} = LDL^T$. This gives the transformed pull-in region

$$F^{-1}(S_{0,\mathsf{B}}) = \{y\in\mathbb{R}^n \mid \mid c_i^T y \mid \leq \frac{1}{2}\ , i = 1,\ldots,n\}$$

Note that $\frac{1}{\mu^2}D$ is a diagonal matrix having the scaled conditional variances $\frac{1}{\mu^2}\sigma^2_{i|I}$ as its entries, and the transformed pull-in region has become an origin-centered n-dimensional cube with all side lengths equal to 1. The multivariate integral can therefore be written as a product of one-dimensional integrals:

$$\begin{aligned}
P_f &= \sum_{z\in\mathbb{Z}^n\setminus\{0\}} \prod_{i=1}^{n} \int_{|y_i|\leq\frac{1}{2}} \frac{1}{\frac{1}{\mu}\sigma_{i|I}\sqrt{2\pi}} \exp\{-\frac{1}{2}\left(\frac{y_i+\frac{1}{\mu}c_i^T L^{-1}z}{\frac{1}{\mu}\sigma_{i|I}}\right)^2\}dy \\
&= \sum_{z\in Z^n\setminus\{0\}} \prod_{i=1}^{n} \int_{-\frac{\mu-2c_i^T L^{-1}z}{2\sigma_{i|I}}}^{\frac{\mu+2c_i^T L^{-1}z}{2\sigma_{i|I}}} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}v^2\}dv \\
&= \sum_{z\in\mathbb{Z}^n\setminus\{0\}} \prod_{i=1}^{n} \left(\Phi\left(\frac{\mu-2c_i^T L^{-1}z}{2\sigma_{i|I}}\right)+\Phi\left(\frac{\mu+2c_i^T L^{-1}z}{2\sigma_{i|I}}\right)-1\right)
\end{aligned} \tag{5.28}$$

In a similar way it can be shown that the IAB success rate is given as:

$$P_s = \prod_{i=1}^{n}\left(2\Phi\left(\frac{\mu}{2\sigma_{i|I}}\right)-1\right) \tag{5.29}$$

2-D example

Figure 5.12 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

5.4.2 Integer Aperture Least-Squares

The IA Least-Squares (IALS) estimator can be defined in a similar way as the IAB estimator, cf. (Teunissen 2004b). The aperture pull-in region is then defined as:

$$\Omega_{z,\mathsf{LS}} = \mu S_{z,\mathsf{LS}} \tag{5.30}$$

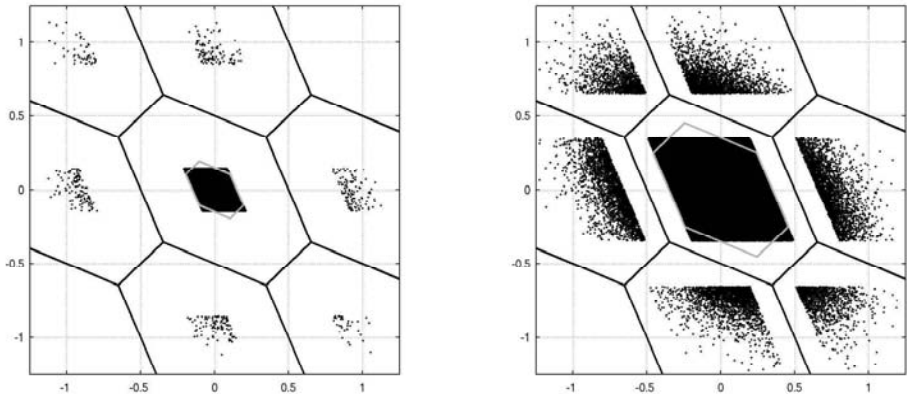

Figure 5.12: 2-D example for IAB estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

with

$$\begin{cases} \mu S_{z,\text{LS}} = \{x \in \mathbb{R}^n \mid \frac{1}{\mu}(x - z) \in S_{0,\text{LS}}\} \\ S_{0,\text{LS}} = \{x \in \mathbb{R}^n \mid \|x\|^2_{Q_{\hat{a}}} \leq \|x - u\|^2_{Q_{\hat{a}}},\ \forall u \in \mathbb{Z}^n\} \end{cases} \tag{5.31}$$

The aperture parameter is μ, $0 \leq \mu \leq 1$.

The first step for computing the IALS estimate is now to compute the ILS solution $\check{a}$ given the float solution $\hat{a}$, and then the ambiguity residuals are up-scaled to $\frac{1}{\mu}\check{\epsilon}$. This up-scaled residual vector is used for verification of:

$$u = \arg\min_{z \in \mathbb{Z}^n} \|\frac{1}{\mu}\check{\epsilon} - z\|^2_{Q_{\hat{a}}} \tag{5.32}$$

If u equals the zero vector, the float solution resides in the aperture pull-in region $\Omega_{0,\text{LS}}$, and thus $\bar{a} = \check{a}$, otherwise $\bar{a} = \hat{a}$.

Since the ILS estimator is optimal, the IALS estimator can be expected to be a better choice than the IAB estimator. The IALS estimate can be computed using the LAMBDA method, and, as for the IAB estimator, the shape of the aperture pull-in region is independent of the chosen aperture parameter. Furthermore, in the Gaussian case it is known that the aperture parameter acts as a scale factor on the vc-matrix, since

$$\begin{aligned} P_s &= \int_{\mu S_{0,\text{LS}}} f_{\hat{a}}(x + a)dx \\ &= \mu^n \int_{S_{0,\text{LS}}} f_{\hat{a}}(\mu x + a)dx \\ &= \int_{S_{0,\text{LS}}} \frac{1}{\sqrt{|\frac{1}{\mu^2}Q_{\hat{a}}|}(2\pi)^{\frac{n}{2}}} \exp\{-\frac{1}{2}\|x\|^2_{\frac{1}{\mu^2}Q_{\hat{a}}}\} \end{aligned}$$

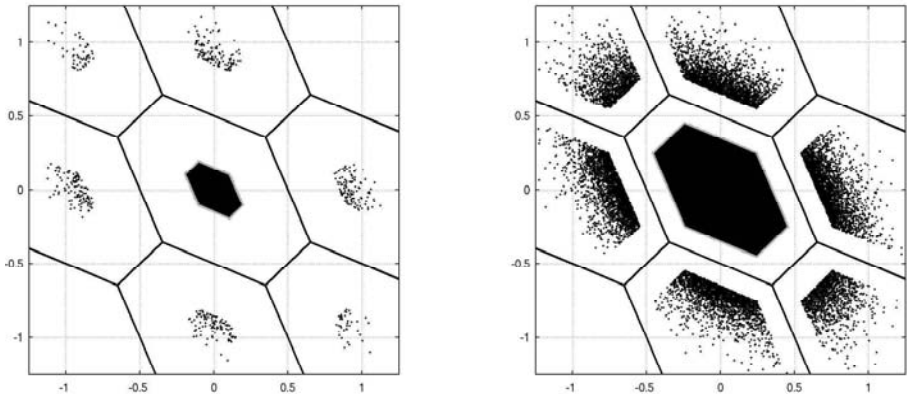

Figure 5.13: 2-D example for IALS estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$.

This shows that the IALS success rate can be computed in the same way as the ILS success rate by replacing $Q_{\hat{a}}$ with the up-scaled version $\frac{1}{\mu^2}Q_{\hat{a}}$. As explained in section 3.2.2 exact evaluation of the ILS success rate is too complex, but the up-scaled version of the vc-matrix can also be used to compute the lower and upper bounds of the IALS success rate as given in section 3.2.2.

Lower and upper bounds for the IALS fail rate can be obtained by bounding the integration region with ellipsoidal regions. The lower bound can be obtained identically to the one for EIA estimation as described in section 5.2. The upper bound follows in a similar way by using $\mu S_0 \subset C_0^\epsilon$ with C_0^ϵ an ellipsoidal region centered at 0 and with size governed by $\epsilon = \mu \max_{x \in S_0} \parallel x \parallel_{Q_{\hat{a}}}$.

Summarizing, the following lower and upper bounds for the IALS success rate and fail rate can be used:

$$P_f \geq \sum_{z \in Z^n \setminus \{0\}} P\left(\chi^2(n, \lambda_z) \leq \frac{1}{4}\mu^2 \min_{z \in \mathbb{Z}^n \setminus \{0\}} \parallel z \parallel^2_{Q_{\hat{a}}}\right) \tag{5.33}$$

$$P_f \leq \sum_{z \in Z^n \setminus \{0\}} P\left(\chi^2(n, \lambda_z) \leq \mu^2 \max_{x \in S_0} \parallel x \parallel^2_{Q_{\hat{a}}}\right) \tag{5.34}$$

$$P_s \geq \prod_{i=1}^{n} [2\Phi(\frac{\mu}{2\sigma_{i|I}}) - 1] \tag{5.35}$$

$$P_s \leq P\left(\chi^2(n, 0) \leq \frac{\mu^2 c_n}{ADOP^2}\right) \tag{5.36}$$

with $\lambda_z = z^T Q_{\hat{a}}^{-1} z$, and c_n and ADOP as given in section 3.2.2. Note that the lower bound of the success rate is equal to the IAB success rate.

If IALS estimation with a fixed fail rate is applied, especially the upper bound of the fail rate is interesting. Unfortunately, equation (5.34) is not useful then, since $\max_{x \in S_0} \parallel x \parallel^2_{Q_{\hat{a}}}$ must be known.

2-D example

Figure 5.13 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

5.5 Penalized Integer Aperture estimation

Penalized Integer Aperture estimation (PIA) was introduced in Teunissen (2004e), and is based on the idea to assign a penalty to each of the possible outcomes of a decision process, and then to minimize the average penalty. In the case of the class of integer aperture estimators defined in equation (5.1) there are three outcomes and the corresponding penalties are denoted as p_s for success, p_f for failure, and p_u for undecided. The expectation of the penalty is given by:

$$E\{p\} = p_s P_s + p_f P_f + p_u P_u \tag{5.37}$$

So, the average penalty is a weighted sum of the individual penalties, and depends on the chosen penalties, and on the aperture pull-in region Ω_0 since this region determines the probabilities P_s, P_f and P_u as can be seen in equation (5.4). The goal is now to find the Ω that minimizes this average penalty, i.e.

$$\min_{\Omega} E\{p\} \tag{5.38}$$

This implies that the aperture space depends on the penalties. The solution is obtained by minimization of:

$$\begin{aligned} E\{p\} &= p_s \int_{\Omega_0} f_{\hat{a}}(x+a)dx + p_f \left(\int_{\Omega_0} (f_{\check{\epsilon}}(x) - f_{\hat{a}}(x+a))dx \right) + p_u \left(1 - \int_{\Omega_0} f_{\check{\epsilon}}(x)dx \right) \\ &= (p_s - p_f) \int_{\Omega_0} f_{\hat{a}}(x+a)dx + (p_f - p_u) \left(\int_{\Omega_0} f_{\check{\epsilon}}(x)dx \right) + p_u \end{aligned} \tag{5.39}$$

For the minimization of $E\{p\}$ with respect to Ω only the first two terms on the right-hand side of equation (5.39) need to be considered, i.e. Ω_0 follows from the following minimization problem:

$$\min_{\Omega_0} \int_{\Omega_0} \underbrace{((p_f - p_u) f_{\check{\epsilon}}(x) + (p_s - p_f) f_{\hat{a}}(x+a))}_{F(x)} dx \tag{5.40}$$

It follows thus that Ω_0 must be chosen such that **all** function values of $F(x)$ are negative:

$$\Omega_{0,\text{PIA}} = \{x \in S_0 \mid f_{\check{\epsilon}}(x) \leq \underbrace{\frac{p_f - p_s}{p_f - p_u}}_{\mu} f_{\hat{a}}(x+a)\} \tag{5.41}$$

Note that the aperture space still depends on the integer estimator that is used, since the pull-in region S_0 is different for rounding, bootstrapping and integer least-squares.

Therefore, the average penalty also has to be minimized for S_0. For that purpose the average penalty is written as:

$$\begin{aligned} E\{p\} &= p_u + (p_s - p_u)(P_s + P_f) + (p_f - p_s)P_f \\ &= p_u + (p_s - p_u)\int\limits_{\Omega} f_{\hat{a}}(x)dx + (p_f - p_s)\sum_{z\in\mathbb{Z}^n\setminus\{a\}}\int\limits_{\Omega\cap S_0} f_{\hat{a}}(x+z)dx \end{aligned} \tag{5.42}$$

Hence, for this minimization problem, only the third term in equation (5.42) needs to be considered. From the optimality of the ILS estimator, cf. (Teunissen 1999b), it follows that

$$\int\limits_{S_{0,\mathsf{LS}}} f_{\hat{a}}(x+a)dx \geq \int\limits_{S_0'} f_{\hat{a}}(x+a)dx \tag{5.43}$$

for any $S_0^{'}$ that fulfills the conditions in (3.1.1). Therefore also the following is true:

$$\int\limits_{\Omega\cap S_{0,\mathsf{LS}}} f_{\hat{a}}(x+a)dx \geq \int\limits_{\Omega\cap S_0'} f_{\hat{a}}(x+a)dx \tag{5.44}$$

This result is used in the identity

$$\int\limits_{\Omega} f_{\hat{a}}(x+a)dx = \sum_{z\in Z^n}\int\limits_{\Omega\cap S_{z,\mathsf{LS}}} f_{\hat{a}}(x+a)dx = \sum_{z\in Z^n}\int\limits_{\Omega\cap S_z'} f_{\hat{a}}(x+a)dx$$

so that

$$\sum_{z\in\mathbb{Z}^n\setminus\{0\}}\int\limits_{\Omega\cap S_{z,\mathsf{LS}}} f_{\hat{a}}(x+a)dx \leq \sum_{z\in\mathbb{Z}^n\setminus\{0\}}\int\limits_{\Omega\cap S_z'} f_{\hat{a}}(x+a)dx \tag{5.45}$$

and thus

$$\int\limits_{\Omega\cap S_{0,\mathsf{LS}}}\sum_{z\in\mathbb{Z}^n\setminus\{0\}} f_{\hat{a}}(x+a+z)dx \leq \int\limits_{\Omega\cap S_0'}\sum_{z\in\mathbb{Z}^n\setminus\{0\}} f_{\hat{a}}(x+a+z)dx \tag{5.46}$$

This shows that the minimization problem is solved if S_0 is chosen as:

$$S_{0,\mathsf{PIA}} = \{x\in\mathbb{R}^n \mid 0 = \arg\max_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+a+z)\} = S_{0,\mathsf{LS}} \tag{5.47}$$

since (5.47) is independent of Ω, and thus (5.38) is indeed solved by choosing Ω_0 and S_0 as in equations (5.41) and (5.47) respectively.

The steps to arrive at the PIA solution are as follows. First the ILS solution $\check{a}$ is computed given the float solution $\hat{a}$. Then the ambiguity residuals $\check{\epsilon}_{\mathsf{LS}}$ are used to verify whether or not

$$\frac{f_{\check{\epsilon}}(\check{\epsilon}_{\mathsf{LS}})}{f_{\hat{a}}(\check{\epsilon}_{\mathsf{LS}}+a)} \leq \mu \tag{5.48}$$

If the inequality holds, $\check{\epsilon}_{\mathsf{LS}} \in \Omega_{0,\mathsf{PIA}}$, and $\bar{a} = \check{a}$, otherwise $\bar{a} = \hat{a}$.

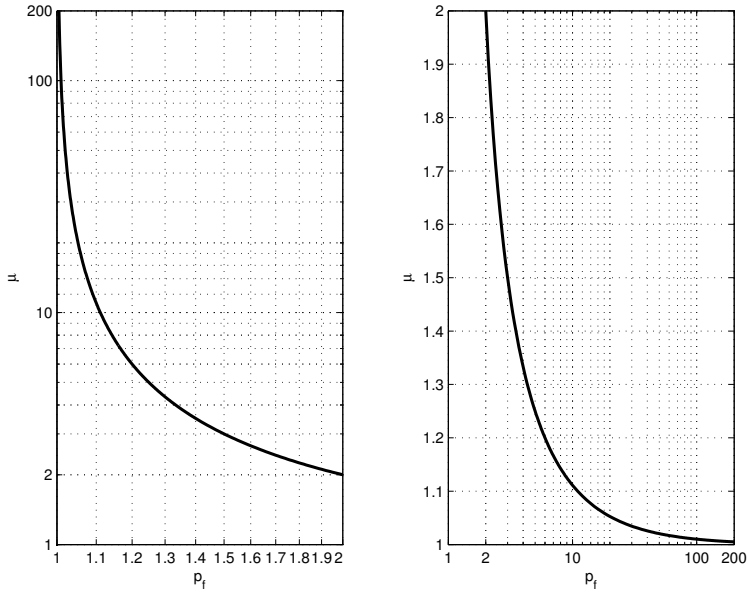

Figure 5.14: μ as function of the penalty p_f with $p_s = 0$ and $p_u = 1$. Left: $p_f = 1, \ldots, 2$; Right: $p_f = 2, \ldots, 200$.

2-D example

In order to get an idea of the shape of the aperture space, $Q_{\hat{z}\_02\_01}$ from appendix B is chosen as the vc-matrix, the corresponding PDFs of the float ambiguities and the ambiguity residuals are determined, and then the inequality (5.48) is used to find Ω_0 as function of μ.

Figure 5.15 shows the results. Furthermore, figure 5.14 shows μ as function of the penalty $p_f > 1$, where $p_s = 0$ and $p_u = 1$. In this case $\mu \to 1$ for $p_f \downarrow p_u$. The reasoning behind this choice is that it seems realistic that no penalty is assigned in the case of correct integer ambiguity resolution, and also that the penalty in the case of incorrect integer ambiguity estimation is larger than the penalty in the case of 'no decision' (undecided). This implies that in order to minimize the average penalty $E\{p\}$, the fail rate P_f should be as small as possible in all cases, and the probability of no decision P_u should also be small, especially if $p_f \downarrow p_u$. Thus the success rate should be as large as possible in all cases, which is the case if the PDF of $\hat{a}$ is peaked. However, for very small values of μ (close to one), the undecided rate will become large since then the aperture space will be small. This results in lower success rates.

It can be seen that the shape of Ω_0 is a mixture of an ellipse and the shape of the pull-in region. The reason is that the contour lines of $f_{\hat{a}}(x)$ (see top panel of figure 5.15) are ellipse-shaped, but those of $f_{\hat{\epsilon}}(x)$ (center panels) are not. It is clear that for larger μ the aperture space becomes larger.

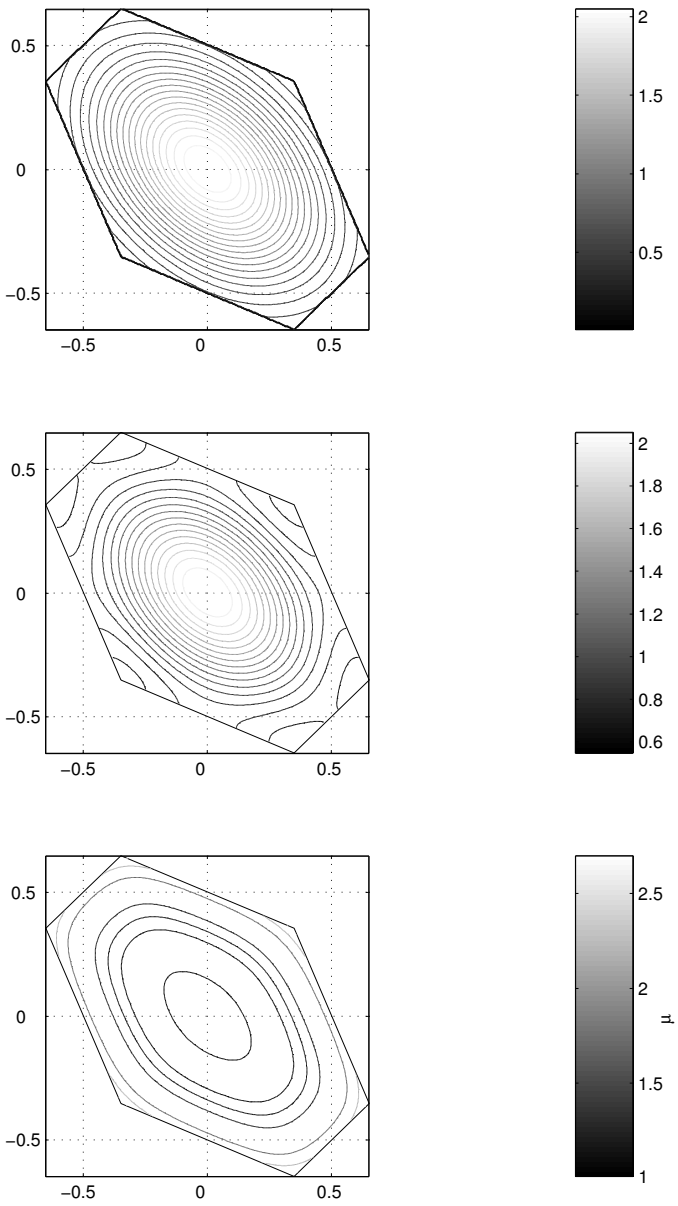

Figure 5.15: Aperture space Ω_0 as function of the penalties, for $Q_{\hat{z}\_02\_01}$. Top: PDF of $\hat{a}$ (within pull-in region); Center: PDF of $\check{\epsilon}$; Bottom: Ω_0 as function of μ.

5.6 Optimal Integer Aperture estimation

The method of penalized IA estimation as described in section 5.5 is some way regulates the probabilities of success and failure. As mentioned in the beginning of this chapter, this is why the new class of integer estimators was defined. A disadvantage of PIA estimation is that the user has to choose the values for the penalties that will be assigned to the three cases success, failure and undecided. A user would like to choose the penalties in such a way that he can be sure that the fail rate is below a certain value. All other IA estimators described in previous sections are not optimal in the sense of maximization of the success rate. Therefore, still another approach could be based on the optimization problem defined as:

$$\max_{\Omega_0 \subset S_0} P_s \quad \text{subject to:} \quad P_f = \beta \tag{5.49}$$

where β is a chosen value for the fail rate, (Teunissen 2003g). In words, this means that the aperture space is defined such that the success rate is maximized subject to a fixed fail rate. The solution of the optimization problem is given by:

$$\Omega_0 = \{x \in S_0 \mid \sum_{z \in \mathbb{Z}^n} f_{\hat{a}}(x+z) \leq \mu f_{\hat{a}}(x+a)\} \tag{5.50}$$

The proof was given in Teunissen (2004d), and is as follows.

Lemma 5.6.1 Let $f(x)$ and $g(x)$ be integrable functions over $\mathbb{R}^n$. Then the region

$$\hat{\Omega} = \{x \in \mathbb{R}^n \mid \sum_{z \in \mathbb{Z}^n} f(x+z) \geq \lambda \sum_{z \in \mathbb{Z}^n} g(x+z),\ \lambda \in \mathbb{R}\} \tag{5.51}$$

solves the constrained maximization problem:

$$\max_{\Omega \subset \mathbb{R}^n} \int_{\Omega} f(x)dx \tag{5.52}$$

subject to:

$$\begin{aligned} &\int_{\Omega \subset \mathbb{R}^n} g(x)dx = c \\ &\Omega = \Omega + z,\ \forall z \in \mathbb{Z}^n \end{aligned} \tag{5.53}$$

if $\lambda \in \mathbb{R}$ is chosen such that the integral constraint of equation (5.53) is satisfied.

Note the similarity with the Neyman-Pearson lemma, cf. (Rao 1965). The proof of lemma 5.6.1 is as follows.

Clearly, the region $\hat{\Omega}$ in (5.51) satisfies the constraints in (5.53). Let Ω be any region that also satisfies both constraints. Then:

$$\int_{\hat{\Omega}} f(x)dx - \int_{\Omega} f(x)dx =$$

$$= \sum_{z\in\mathbb{Z}^n} \int_{\hat{\Omega}\cap S_z} f(x)dx - \sum_{z\in\mathbb{Z}^n} \int_{\Omega\cap S_z} f(x)dx \quad \text{(a)}$$

$$= \int_{\hat{\Omega}\cap S_0} \sum_{z\in\mathbb{Z}^n} f(x+z)dx - \int_{\Omega\cap S_0} \sum_{z\in\mathbb{Z}^n} f(x+z)dx \quad \text{(b)}$$

$$= \int_{(\hat{\Omega}\setminus\{\Omega\})\cap S_0} \sum_{z\in\mathbb{Z}^n} f(x+z)dx - \int_{(\Omega\setminus\{\hat{\Omega}\})\cap S_0} \sum_{z\in\mathbb{Z}^n} f(x+z)dx \quad \text{(c)}$$

$$\geq \lambda \int_{(\hat{\Omega}\setminus\{\Omega\})\cap S_0} \sum_{z\in\mathbb{Z}^n} g(x+z)dx - \lambda \int_{(\Omega\setminus\{\hat{\Omega}\})\cap S_0} \sum_{z\in\mathbb{Z}^n} g(x+z)dx \quad \text{(d)}$$

$$= \lambda \int_{\hat{\Omega}\setminus\{\Omega\}} g(x)dx - \lambda \int_{\Omega\setminus\{\hat{\Omega}\}} g(x)dx = 0 \Longrightarrow \quad \text{(e)}$$

$$\int_{\hat{\Omega}} f(x)dx \geq \int_{\Omega} f(x)dx$$

where the following properties were used:

(a) $\bigcup_{z\in\mathbb{Z}^n} S_z = \mathbb{R}^n$

(b) $S_z = S_0 + z$

(c) $(\hat{\Omega}\setminus\{\Omega\}) \cap S_0 = (\hat{\Omega} - \hat{\Omega}\cap\Omega) \cap S_0$ common part of $\hat{\Omega}$ and Ω subtracted

(d) use (5.51)

(e) $S_z = S_0 + z$ and use (5.53)

Recall that the success rate and fail rate can be expressed as:

$$P_s = \int_{\Omega_0} f_{\hat{a}}(x+a)dx, \quad P_f = \sum_{z\in\mathbb{Z}^n\setminus\{a\}} \int_{\Omega_0} f_{\hat{a}}(x+z)dx$$

so that

$$P_s + P_f = \int_{\Omega} f_{\hat{a}}(x)dx$$

$$P_f = \int_{\Omega\setminus\{\Omega_a\}} f_{\hat{a}}(x)dx = \int_{\Omega} f_{\hat{a}}(x)(1 - s_a(x))dx$$

The optimization problem of (5.49) is thus equivalent to:

$$\max_{\Omega_0}(P_s + P_f) \quad \text{subject to:} \quad P_f = \beta \tag{5.54}$$

or to

$$\max_{\Omega} \int_{\Omega} f_{\hat{a}}(x)dx \quad \text{subject to:} \quad \int_{\Omega} f_{\hat{a}}(x)(1-s_a(x))dx = \beta,\ \Omega = \Omega + z,\ z \in \mathbb{Z}^n \tag{5.55}$$

Lemma 5.6.1 can now be applied to obtain the following result:

$$\Omega = \{x \in \mathbb{R}^n \mid \sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z) \geq \lambda \sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z)(1-s_a(x+z))\}$$

Therefore

$$\begin{aligned}\Omega_0 &= \Omega \cap S_0 \\ &= \{x \in \mathbb{R}^n \mid \sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z) \geq \lambda(\sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z) - f_{\hat{a}}(x+a)),\ x \in S_0\} \\ &= \{x \in \mathbb{R}^n \mid \sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z) \leq \frac{\lambda}{\lambda-1} f_{\hat{a}}(x+a),\ x \in S_0\}\end{aligned}$$

This ends the proof of (5.50).

The best choice for S_0 is the ILS pull-in region. The reason is that $P_s + P_f$ is independent of S_0, but P_s and P_f are not. Therefore, any choice of S_0 which makes P_s larger, automatically makes P_f smaller. So indeed the best choice for S_0 follows from:

$$\max_{S_0} \int_{\Omega\cap S_0} f_{\hat{a}}(x+a)dx \quad \text{subject to} \quad \Omega = \Omega + z,\ \forall z \in \mathbb{Z}^n$$

as the ILS pull-in region, see (Teunissen 1999a).

Recall that $\sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(x+z)s_0(x) = f_{\check{\epsilon}}(x)$. It follows thus that the aperture space is defined in the same way as for PIA estimation, only that μ does not depend on the penalties anymore, but on the choice of $P_f = \beta$. This new approach will be referred to as Optimal IA estimation (OIA estimation).

In practice it will be difficult to compute μ for a certain choice of the fail rate. Instead, one could use the following upper bound:

$$\begin{aligned}\beta = P_f &= \sum_{z\in\mathbb{Z}^n\setminus\{a\}} \int_{\Omega_0} f_{\hat{a}}(x+z)dx \\ &\leq \int_{\Omega_0} (\mu-1) f_{\hat{a}}(x+a)dx = (\mu-1)\int_{\Omega_a} f_{\hat{a}}(x)dx \\ &\leq (\mu-1)\int_{S_a} f_{\hat{a}}(x)dx \leq (\mu-1)P_{s,\text{LS}}\end{aligned} \tag{5.56}$$

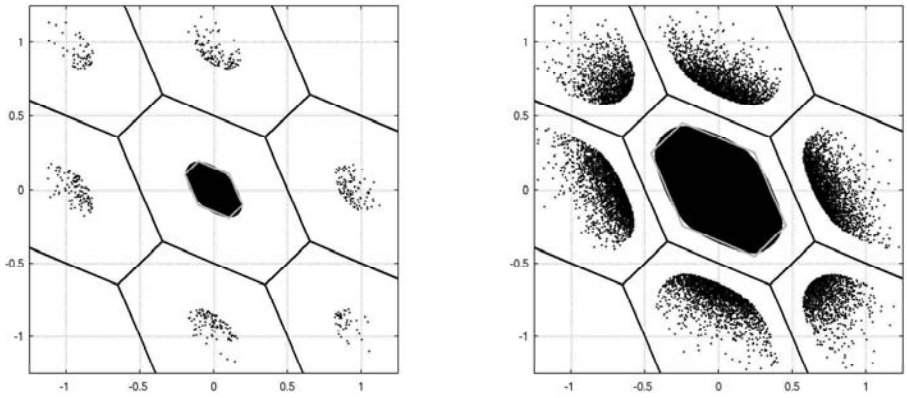

Figure 5.16: 2-D example for OIA estimation. Float samples for which $\bar{a} = \check{a}$ are shown as black dots. Left: $P_f = 0.001$; Right: $P_f = 0.025$. Also shown is the IA least-squares pull-in region for the same fail rates (in grey).

Finally, one could use the upper bound of the ILS success rate as defined in (3.39) in order to find a value for μ so that the fail rate will be smaller than or equal to $(\mu-1)P_{s,\text{LS}}$. However, it can be expected that this upper bound is not tight enough, resulting in an aperture space that is much smaller than actually needed, see the example below.

2-D example

In order to evaluate the accuracy of equation (5.56), the success and fail rates for two values of μ were determined based on the simulations. The following was obtained:

$$\begin{aligned} \mu = 1.01: &\quad P_f = 0.0010 \leq (\mu - 1)P_s = 0.0016 \\ \mu = 1.25: &\quad P_f = 0.0386 \leq (\mu - 1)P_s = 0.1773 \end{aligned}$$

which shows that already $(\mu - 1)P_s$ is not a strict upper bound. So, it will certainly not make sense to use $(\mu - 1)P_{s,\text{LS}}$ as an upper bound for the fail rate.

Figure 5.16 shows in black all float samples for which $\bar{a} = \check{a}$ for two different fail rates.

The results obtained with all IA estimators are summarized in table 5.1; the penalized IA estimator is not considered, since it is identical to the optimal IA estimator if the fixed fail rate approach is applied. For the two fixed fail rates, the corresponding aperture parameters and the success rates are given. Furthermore, the percentage of solutions identical to the OIA solution is given. It follows that especially the RTIA estimator performs almost equally well as the OIA estimator. But for this example also EIA, DTIA and IALS estimation perform good. In section 5.8 the different IA estimators will be compared more extensively using geometry-based models.

Table 5.1: Comparison of IA estimators for 2-D example. Two fixed fail rates P_f are used; μ is the aperture parameter, P_s the success rate, and P_{id} is the percentage of identical solutions to the OIA solution.

P_f		EIA	RTIA	F-RTIA	DTIA	PTIA	IAB	IALS	OIA
0.001	μ	0.605	0.035	0.068	10.322	0.540	0.293	0.300	1.011
	P_s	0.168	0.169	0.133	0.165	0.166	0.161	0.169	0.169
	P_{id}	98.4	98.9	89.4	97.6	98.0	96.9	98.7	100
0.025	μ	1.414	0.314	0.386	4.432	1.336	0.690	0.700	1.147
	P_s	0.632	0.634	0.611	0.632	0.629	0.615	0.633	0.634
	P_{id}	97.5	99.2	92.5	97.5	96.5	92.7	97.7	100

5.7 Implementation aspects

The optimal IA estimator was defined such that the aperture space is constructed based on the conditions that the fail rate is equal to a fixed value and at the same time the maximum success rate is obtained. All other IA estimators can be used in the same way: choose a fixed fail rate and determine the aperture space for which this fixed value of the fail rate is obtained. However, these IA estimators are not optimal since the aperture space is not defined such that the maximum success rate is obtained.

The problem with the approach of choosing a fixed fail rate, is the determination of the aperture parameter. Only for the EIA estimator and the IAB estimator exact evaluation of the fail rate as function of the aperture parameter is possible. For all other estimators it will be necessary to determine the aperture parameter empirically or by using simulations. Note that in practice often empirically determined critical values are used for the ratio test and the difference test.

In this section it is first explained how the aperture parameter can be determined for the EIA and IAB estimator. Then a method will be presented which uses simulations in order to determine the aperture parameter. Finally, a procedure will be outlined which can be used for applying IA estimation.

5.7.1 Determination of the aperture parameter using root finding methods

The fail rate in the case of EIA and IAB estimation can be evaluated exactly, see equations (5.12) and (5.28) respectively. The problem is now to determine the aperture parameter such that a fixed fail rate is obtained. For that purpose root finding methods must be used, see appendix A.3. The procedure is then as follows. Choose the fixed fail rate $P_f = \beta$ and then find the root of

$$f(\mu) = P_f(\mu) - \beta$$

using one of the root finding methods.

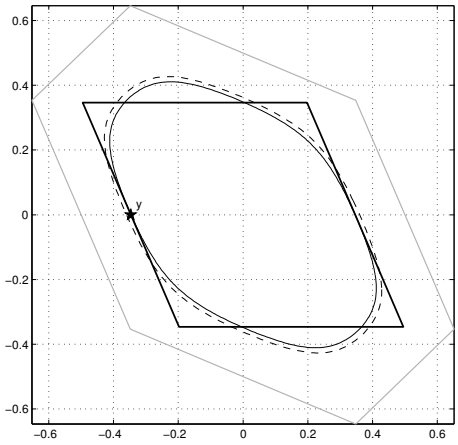

Figure 5.17: Aperture parameter for OIA estimation approximated using IAB estimation with $P_{f,\text{IAB}} = 0.025$. y is an element of the boundary of both the IAB pull-in region and the approximated OIA pull-in region. The dashed OIA pull-in region is the one for which $P_{f,\text{OIA}} = 0.025$.

Since it is quite easy to determine the aperture parameter μ for a fixed fail rate in the case of IAB or EIA estimation, it would be interesting to investigate whether this parameter can be used to approximate the aperture parameter for e.g. Optimal IA estimation such that $P_{f,\text{OIA}}(\mu') \approx P_{f,\text{IAB}}(\mu)$ with μ' the approximated aperture parameter.

One approach would be to determine the aperture parameter for IAB estimation using a root finding method. Then an ambiguity vector on the boundary of the bootstrapped aperture pull-in region $\Omega_{0,\text{B}}$ is determined. From the definition of the bootstrapped pull-in region in equation (3.12) it can be seen that $S_{0,\text{B}}$ is constructed as the intersection of half-spaces which always pass through the points $\pm\frac{1}{2}c_i$, with c_i the canonical vector with a one as its ith entry. From equation (3.10) follows then that $\hat{a} = \frac{1}{2}c_n$ will lie on the boundary of $S_{0,\text{B}}$ and $S_{c_n,\text{B}}$ since

$$\check{a}_{j,\text{B}} = \left[\hat{a}_j - \sum_{i=1}^{j-1}\sigma_{\hat{a}_j\hat{a}_{i|I}}\sigma^{-2}_{\hat{a}_{i|I}}(\hat{a}_{i|I} - \check{a}_i)\right] = 0, \quad \forall j = 1,\ldots,n-1$$

$$\check{a}_{n,\text{B}} = \left[\hat{a}_n - \sum_{i=1}^{n-1}\sigma_{\hat{a}_n\hat{a}_{i|I}}\sigma^{-2}_{\hat{a}_{i|I}}\underbrace{(\hat{a}_{i|I} - \check{a}_i)}_{=0\,\forall i}\right] = [\hat{a}_n] = \left[\frac{1}{2}\right]$$

$$\Longrightarrow \begin{cases}\check{a}_\text{B} = 0 & \text{or}\\ \check{a}_\text{B} = c_n\end{cases}$$

Now y is chosen as:

$$y = \mu \cdot \frac{1}{2}c_n, \tag{5.57}$$

since this will be an element of the boundary of the scaled bootstrapped pull-in region, i.e. the IAB pull-in region. It must then be checked if $y \in S_{0,\text{LS}}$, which will generally

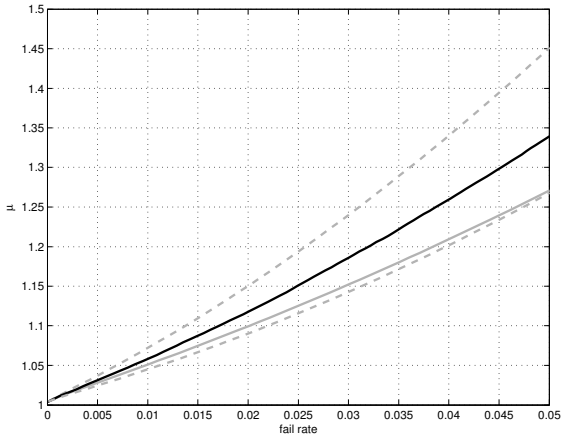

Figure 5.18: Aperture parameter for OIA estimation as function of the fail rate. Black: determined with simulations; Grey solid: approximated using IAB estimation; Grey dashed: approximated using EIA estimation.

be the case when the ambiguities are decorrelated, since then the bootstrapped and the ILS pull-in regions largely overlap. If that is the case the aperture parameter for OIA estimation is approximated as:

$$\mu' = \frac{f_{\check{\epsilon}}(y)}{f_{\hat{a}}(y)}, \tag{5.58}$$

since then y is also an element of the boundary of $\Omega_{0,\text{OIA}}$. If the aperture pull-in regions of IAB and OIA would have a similar shape, this approximation may work well. However, it does not tell whether using μ' will give a larger or smaller fail rate than the fixed value that was used in order to determine μ'.

This approach is illustrated for the 2-D example in figure 5.17. The IAB pull-in region for a fail rate of 0.025 is shown. The star shows y as determined with equation (5.57)[1]. The resulting OIA pull-in region is shown with a solid line; y is an element of the boundary of this pull-in region. This 'approximated' OIA pull-in region is smaller than the one for which $P_{f,\text{OIA}} = 0.025$ (dashed line).

Figure 5.18 shows the OIA aperture parameter as function of the fail rate, where the aperture parameter is determined with simulations (see section 5.7.2), or approximated using IAB estimation. It can be seen that the approximated aperture parameter is too low, which means that $P_{f,\text{OIA}}(\mu') < \beta$. In other words, for this example the OIA pull-in regions are always smaller than necessary if the aperture parameter is approximated using IAB estimation.

A similar approach can be followed by using EIA estimation. Then a vector on the

[1]Note that if bootstrapping is applied, one starts with the first ambiguity which should be the most precise float ambiguity as was explained in section 3.1.2. For the example shown here, the second ambiguity is the most precise. Hence, y is permuted so that $y = \mu \cdot \frac{1}{2}c_1$.

boundary of the aperture pull-in region $\Omega_{0,\mathsf{E}}$ could be chosen as:

$$y = \varrho \cdot \frac{1}{2} \arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|^2_{Q_{\hat{a}}} \quad \text{with} \quad \varrho = \frac{\mu}{\frac{1}{2} \min\limits_{z \in \mathbb{Z}^n \setminus \{0\}} \|z\|_{Q_{\hat{a}}}} \tag{5.59}$$

since then $\|y\|^2_{Q_{\hat{a}}} = \mu^2$. Again the aperture parameter for OIA estimation is approximated with equation (5.58), so that y is also an element of the boundary of $\Omega_{0,\mathsf{OIA}}$. It is important to note that this approach will not work in the case the ellipsoidal pull-in regions overlap, since then $y \notin S_0$. In that case ϱ in equation (5.59) should be set equal to one, so that y is on the boundary of S_0.

Figure 5.18 also shows the OIA aperture parameter as function of the fail rate as approximated using EIA. Instead of choosing y as in equation (5.59), one could also choose $y = \varrho \cdot \frac{1}{2} u$ with u the adjacent integer with the largest distance to the true integer $a = 0$. With the first option, the approximated aperture parameter will be too large, with the second option the approximated aperture parameter will be too small. For this example, approximation with IAB estimation works better, but still not good.

5.7.2 Determination of the aperture parameter using simulations

As explained at the beginning of this section, computation of the aperture parameter μ for a fixed fail rate is not possible for most IA estimators. An alternative is to determine μ numerically. Therefore, simulated data are required. The procedure will be described here for the OIA estimator, but a similar approach can be followed for other IA estimators.

The procedure is as follows:

1. Generate N samples of normally distributed float ambiguities:
   $\hat{a}_i \sim N(0, Q_{\hat{a}})$, $i = 1, \ldots, N$

2. Determine the ILS solution:
   $\check{a}_i, \qquad \check{\epsilon}_i = \hat{a}_i - \check{a}_i, \qquad r_i = \frac{f_{\hat{\epsilon}}(\check{\epsilon}_i)}{f_{\hat{a}}(\check{\epsilon}_i)}$

3. Choose the fail rate: $P_f = \beta \leq P_{f,\mathsf{LS}}$
   The fail rate as function of μ is given by: $P_f(\mu) = \frac{N_f}{N}$
   with $N_f = \sum\limits_{i=1}^{N} \omega(r_i, \check{a}_i)$ and $\omega(r_i, \check{a}_{i,}) = \begin{cases} 1 & \text{if } r_i \leq \mu \ \wedge \ \check{a}_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$

4. Choose:
   $\mu_1 = (\min(r_i) - 10^{-16})$ since this results in: $N_f = 0 \Rightarrow P_f(\mu_1) = 0$
   $\mu_2 = \max(r_i)$ since this results in: $P_f(\mu_2) = P_{f,\mathsf{LS}}$

5. Use a root finding method to find $\mu \in [\mu_1, \mu_2]$ so that $P_f(\mu) - \beta = 0$.
   This will give the solution since $P_f(\mu_1) - \beta < 0$ and $P_f(\mu_2) - \beta > 0$, and the fail rate is monotonically increasing for increasing μ, as will be shown below.

In order to show that the fail rate is monotonically increasing with μ choose $\mu_j > \mu_i$. Assume the aperture space $\Omega_{0,i}$ is defined by μ_i, so that the following is true:

$$\frac{f_{\check{\epsilon}}(x)}{f_{\hat{a}}(x)} \leq \mu_i \, \forall x \in \Omega_{0,i}$$

Since $\mu_j > \mu_i$, the following is true:

$$\frac{f_{\check{\epsilon}}(x)}{f_{\hat{a}}(x)} < \mu_j \quad \forall x \in \Omega_{0,i} \quad \text{and} \quad \frac{f_{\check{\epsilon}}(x)}{f_{\hat{a}}(x)} \leq \mu_j \quad \forall x \in \Omega_{0,j}$$

This means that $\Omega_{0,i} \subset \Omega_{0,j}$ and thus also $\Omega_{z,i} \subset \Omega_{z,j}, \forall z \in \mathbb{Z}^n$. From the definition of the fail rate in (5.4) it follows that $P_{f,j} > P_{f,i} \, \forall \mu_j > \mu_i$.

5.7.3 Determination of the Integer Aperture estimate

In practice the problem of determining the aperture parameter for a given fail rate can be circumvented as follows. First, the float and fixed solution are computed, and then it is determined what the fail rate would be if the float solution would fall on the boundary of the aperture pull-in region $\Omega'_{\hat{a}}$. If this fail rate is larger than the user-defined threshold β, obviously Ω' is larger than allowed and the float solution lies outside the aperture space Ω corresponding to $P_f = \beta$. The procedure is outlined below for the RTIA estimator, but can also be applied for other IA estimators. For the IALS estimator this is not completely straightforward, and therefore in appendix D the procedure for this estimator is outlined.

1. Choose fixed fail rate: $P_f = \beta$
2. Collect observations and apply least-squares adjustment: $\hat{a}$, $Q_{\hat{a}}$;
3. Apply integer least-squares estimation: $\check{a}$, $\check{\epsilon}$, $\|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}}$, $\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}$;
4. Choose $\mu' = \frac{\|\hat{a}-\check{a}\|^2_{Q_{\hat{a}}}}{\|\hat{a}-\check{a}_2\|^2_{Q_{\hat{a}}}}$;
5. Generate N samples of float ambiguities $\hat{x}_i \sim N(0, Q_{\hat{a}})$;
6. For each sample determine the ILS estimate $\check{x}_i$
   and $\mu_i = \frac{\|\hat{x}_i-\check{x}_i\|^2_{Q_{\hat{a}}}}{\|\hat{x}_i-\check{x}_{i,2}\|^2_{Q_{\hat{a}}}}$;
7. Count number of samples N_f for which: $\mu_i \leq \mu'$ and $\check{x}_i \neq 0$;
8. Compute fail rate with μ' as critical value: $P_f(\mu') = \frac{N_f}{N}$;
9. If $P_f(\mu') \leq \beta$: $\bar{a} = \check{a}$, otherwise $\bar{a} = \hat{a}$.

In order to get a good approximation of the fail rate in step 8 the number of samples must typically be chosen large, $N > 100,000$. But in order to reduce the computation

Table 5.2: Comparison of IA estimators for geometry-based examples.

	fail rate	success rates							
		EIA	RTIA	F-RTIA	DTIA	PTIA	IAB	IALS	OIA
06_01	0.001	0.158	0.165	0.131	0.163	0.162	0.134	0.166	0.166
06_01	0.008	0.370	0.390	0.365	0.383	0.375	0.316	0.388	0.392
06_02	0.012	0.044	0.044	0.037	0.042	0.042	0.040	0.043	0.044
06_02	0.050	0.129	0.130	0.119	0.124	0.124	0.116	0.127	0.131
10_01	0.001	0.649	0.921	0.923	0.932	0.799	0.846	0.914	0.932
10_01	0.004	0.847	0.968	0.969	0.970	0.892	0.923	0.968	0.970
10_02	0.006	0.111	0.191	0.179	0.208	0.179	0.154	0.176	0.219
10_02	0.012	0.174	0.289	0.279	0.301	0.251	0.236	0.270	0.314
10_03	0.020	0.052	0.074	0.070	0.076	0.071	0.066	0.073	0.084
10_03	0.060	0.117	0.161	0.154	0.160	0.145	0.142	0.159	0.170

time, which is important in real-time applications, it has been investigated how much the results would differ with N much smaller so that the computation time becomes acceptable. It followed that for all models described in appendix B, already with a few thousand samples the results are quite good: same decision as with very large N in more than 99% of the cases.

5.8 Comparison of the different IA estimators

In section 3.5 the integer validation tests as currently used in practice were described. If an admissible integer estimator is used in combination with one of the discrimination tests described in that section, the resulting estimator belongs to the class of integer aperture estimators as was shown in section 5.3 for the ratio test, difference test and projector test.

It is now interesting to compare the performance of all the different IA estimators using the fixed fail rate approach. For that purpose simulations are used. It is then investigated which of the IA estimators perform close to optimal. Furthermore, the OIA estimator is conceptually compared with the ratio test and difference test, and with the IALS estimator.

5.8.1 Examples for the geometry-based case

Simulations were carried out for several geometry-based models, see appendix B.

Table 5.2 shows success rates obtained for two different fixed fail rates for all examples. From the results follows that OIA estimation indeed results in the highest success rates, but often the ratio test and the difference test give almost the same results. In section 5.8.2 these estimators are compared in more detail. Also the IALS estimator performs

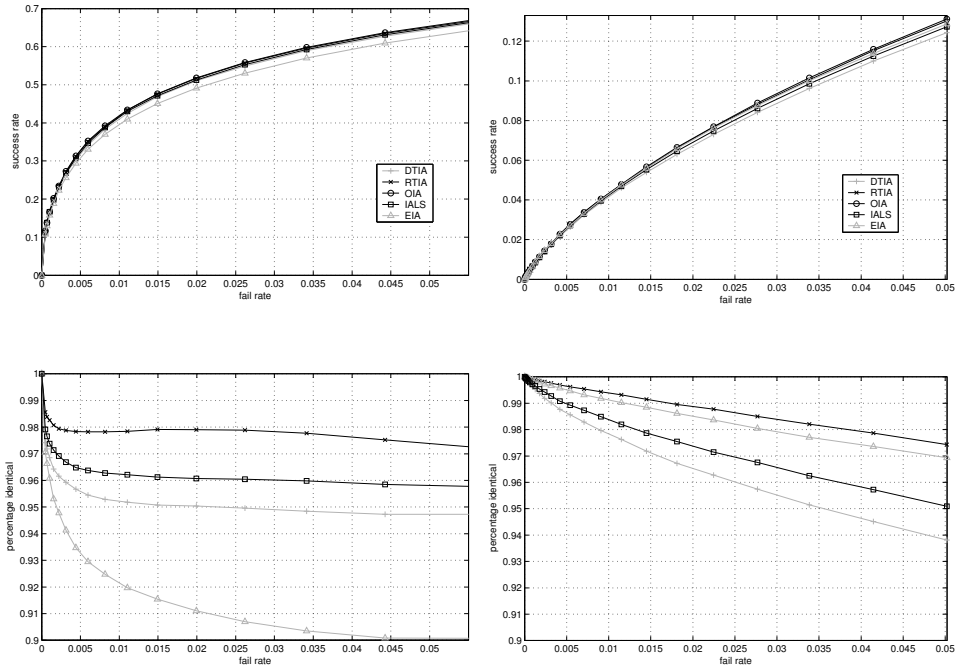

Figure 5.19: Top: Success rate as function of the fail rate; Bottom: percentages of solutions identical to OIA. Left: example 06_01. Right: example 06_02.

quite well. The EIA estimator only performs well for examples 06_01 and 06_02. The performance of the F-RTIA estimator seems to be much worse than the performance of the other estimators. The reason is that the F-RTIA estimator does not only consider the ambiguity residuals, but also the residuals of the float solution, so that the aperture space is not n-dimensional, but $(m-p)$-dimensional.

Figures 5.19 and 5.20 show the success rates as function of the fail rate for those IA estimators that perform reasonably well; only low fail rates are considered. DTIA, RTIA and IALS estimation perform well for all examples: the success rates are almost equal to the OIA success rate, and the percentage of solutions identical to the OIA solution is high. For examples 06_01 and 06_02 the RTIA estimator performs best after OIA estimation; for examples 10_01 and 10_03 the DTIA estimator performs better than RTIA estimation. EIA estimation only performs well for example 06_02.

Until now the performance of the IA estimators is solely based on the ambiguity success rates and fail rates. Of course, a user is more interested in the probability that the baseline estimate will be within a certain distance from the true baseline:

$$P(\|\dot{b} - b\|^2_{Q_{\dot{b}}} \leq \epsilon) \tag{5.60}$$

where $\dot{b}$ is either the float, fixed, or IA baseline estimator. The vc-matrices $Q_{\check{b}}$ and $Q_{\bar{b}}$ can be determined from the estimates corresponding to the simulated float ambiguities. These probabilities are difficult to interpret, but will be considered for completeness, because they tell something about the distributional properties of the baseline estimators.

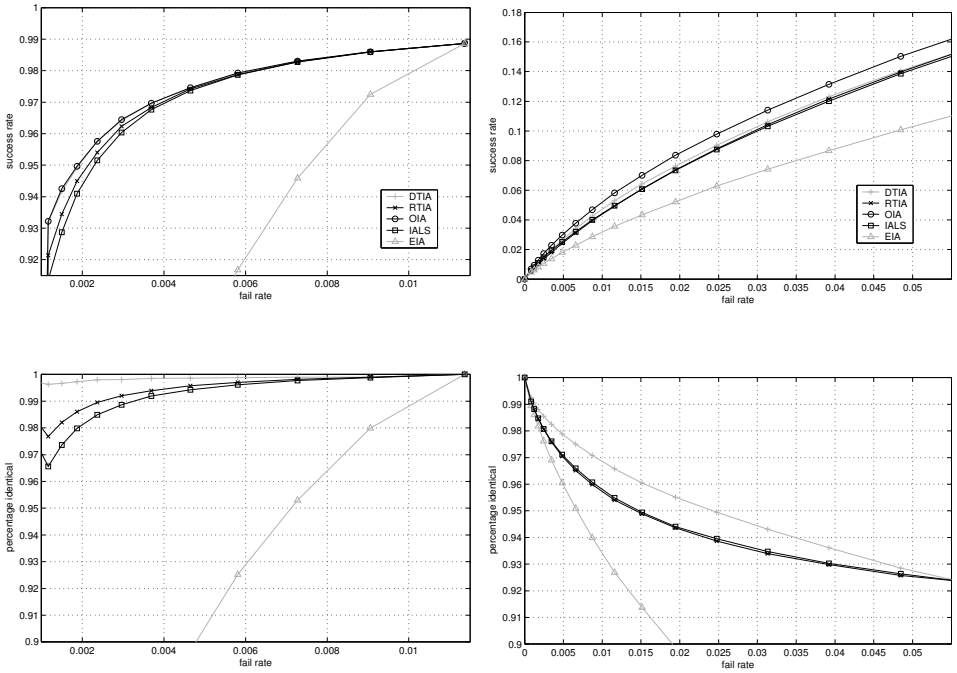

Figure 5.20: Top: Success rate as function of the fail rate; Bottom: percentages of solutions identical to OIA. Left: example 10_01. Right: example 10_03.

A user will be mainly interested in the baseline coordinates, therefore also the following probabilities are analyzed:

$$P(\|\dot{b}_x - b_x\| \leq \epsilon) \tag{5.61}$$

where b_x refers only to the three baseline coordinates. So, the distance to the true position is considered here.

The results are shown in figures 5.21 and 5.22. For all examples the OIA estimates corresponding to the largest fail rate in table 5.2 is used. Note that the same probabilities were analyzed in section 4.3.3 for the float, fixed and BIE baseline estimators.

The probability that $\check{b}$ is close to the true b is highest for most examples. Note, however, that in practice $\check{b}$ should only be used if the ILS success rate is close to one. This is the case for example 10_01, but then the baseline probability of $\bar{b}$ is comparable to that of $\check{b}$. Another disadvantage of the fixed estimator is that if $\check{b}$ is not close to b, there is a high probability that it will be very far from the true solution.

It follows that the baseline probabilities of $\bar{b}$ often are better than those of $\hat{b}$. In the figures only the probabilities of the OIA estimator are shown; for the other IA estimators the probabilities are somewhat lower. If the aperture parameter would be chosen larger or smaller, the OIA baseline probability would be more like that of the fixed or float baseline probability respectively.

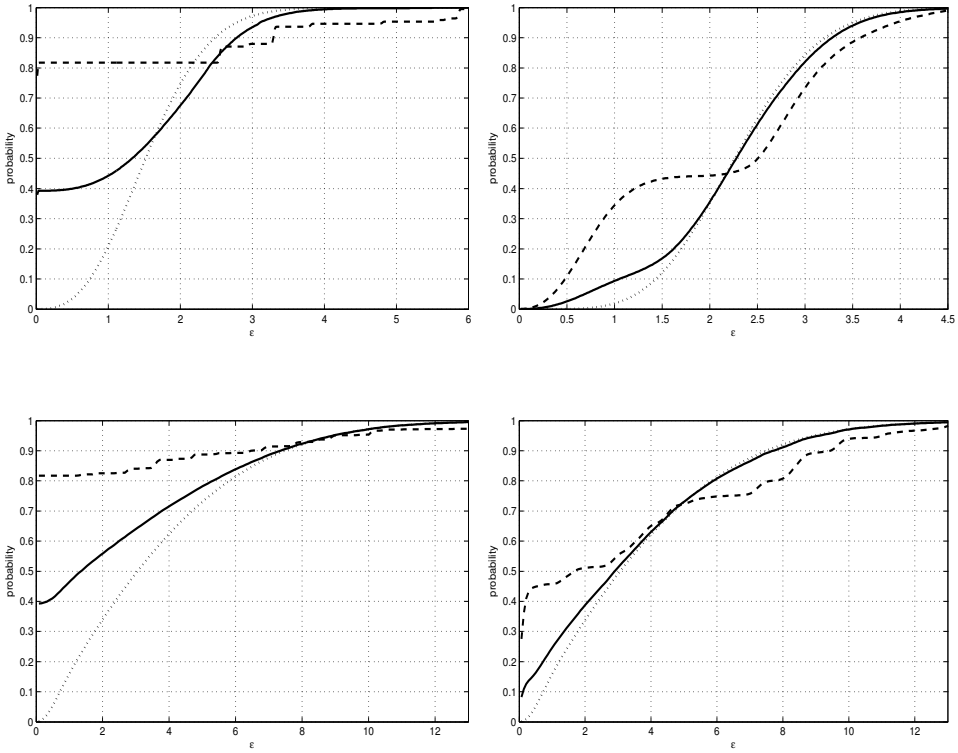

Figure 5.21: Top: $P(\|\dot{b}-b\|^2_{Q_{\dot{b}}} \leq \epsilon)$. Bottom: $P(\|\dot{b}_x - b_x\|^2 \leq \epsilon)$. Dotted: $\dot{b} = \hat{b}$; Dashed: $\dot{b} = \check{b}$; Solid: $\dot{b} = \bar{b}$. Left: example 06_01. Right: example 06_02.

Besides the probabilities as defined in equations (5.60) and (5.61), it is also interesting to know the probability that the IA baseline estimator is better than or identical to the float estimator, i.e.

$$P(\|\bar{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|) \tag{5.62}$$

As explained in section 3.4 the fixed baseline probabilities cannot be evaluated exactly; the same is thus true for the IA estimators. The probability is plotted in figure 5.23 as function of the fail rate for the different examples. The end points correspond to the ILS fail rate and therefore the probability is equal to $P(\|\check{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|)$.

Especially for low fail rates the probabilities are high, and $P(\|\bar{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|) > P(\|\check{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|)$. That is because then either the undecided rate will be high and thus there is a high probability that $\bar{b} = \hat{b}$, or the success rate is high so that there is a high probability that $\check{b}$ is better than $\hat{b}$ if $\bar{b} = \check{b}$. For examples 06_01 and 06_02 the probability for the IA baseline estimator is always better than for the fixed estimator; but that is not true for examples 10_02 and 10_03. Since a user will be interested in low fail rates, these probabilities show that the IA aperture estimator is to be preferred above the fixed estimator.

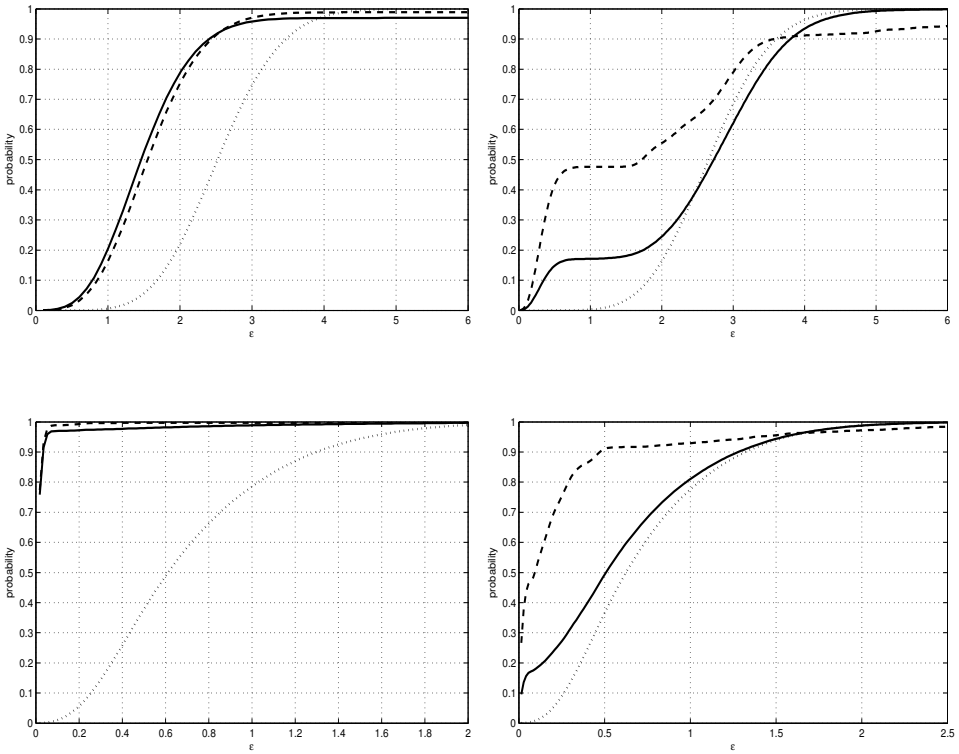

Figure 5.22: Top: $P(\|\dot{b}-b\|^2_{Q_{\dot{b}}} \leq \epsilon)$. Bottom: $P(\|\dot{b}_x - b_x\|^2 \leq \epsilon)$. Dotted: $\dot{b} = \hat{b}$; Dashed: $\dot{b} = \check{b}$; Solid: $\dot{b} = \bar{b}$. Left: example 10_01. Right: example 10_03.

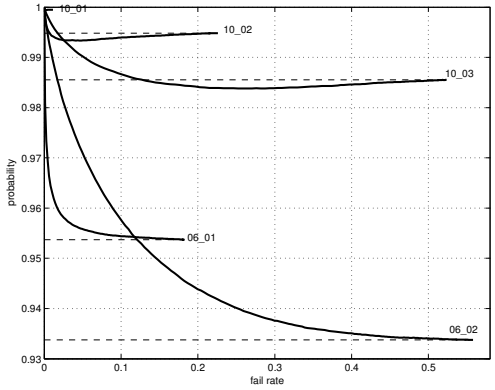

Figure 5.23: Probability $P(\|\bar{b}_x - b_x\| \leq \|\hat{b}_x - b_x\|)$ as function of the fail rate.

5.8.2 OIA estimation versus RTIA and DTIA

The results in the preceding section have shown that the RTIA and the DTIA estimator both perform close to optimal for most examples. Therefore, in this section the relation between these estimators and the OIA estimator is investigated.

Recall that with the ratio test the fixed solution is accepted if and only if:

$$\frac{R_1}{R_2} \leq \mu_{\mathsf{R}}$$

With the difference test the acceptance criterion is:

$$R_2 - R_1 \geq \mu_{\mathsf{D}}$$

With OIA estimation the fixed solution is used if and only if:

$$\frac{f_{\check{\epsilon}}(\hat{a}-\check{a})}{f_{\hat{a}}(\hat{a}-\check{a})} \leq \mu_{\mathsf{O}}$$

with

$$\frac{f_{\check{\epsilon}}(\hat{a}-\check{a})}{f_{\hat{a}}(\hat{a}-\check{a})} = \frac{\sum_{i=1}^{\infty} \exp\{-\frac{1}{2}R_i\}}{\exp\{-\frac{1}{2}R_1\}} = 1 + \frac{\exp\{-\frac{1}{2}R_2\}}{\exp\{-\frac{1}{2}R_1\}} + \frac{\sum_{i=3}^{\infty} \exp\{-\frac{1}{2}R_i\}}{\exp\{-\frac{1}{2}R_1\}} \tag{5.63}$$

where the i refers to the solution that is best (1), second-best (2), etcetera. This expression follows from equations (3.26) and (3.48).

An important difference with the ratio test statistic and the difference test statistic is that in this expression all $R_i, i = 1, \ldots, \infty$ are taken into account and not just those that correspond to the best and second-best integer solution. From equation (5.63) follows that:

$$\frac{f_{\check{\epsilon}}(\hat{a}-\check{a})}{f_{\hat{a}}(\hat{a}-\check{a})} \geq 1 + \exp\{-\frac{1}{2}(R_2 - R_1)\} \tag{5.64}$$

So, if Ω_0' is the aperture pull-in region corresponding to:

$$1 + \exp\{-\frac{1}{2}(R_2 - R_1)\} \leq \mu_{\mathsf{O}} \tag{5.65}$$

if follows that $\Omega_{0,\mathsf{O}} \subset \Omega_0'$. Equation (5.65) can be rewritten as:

$$R_2 - R_1 \geq -2\ln(\mu_{\mathsf{O}} - 1) \tag{5.66}$$

Equation (5.66) is identical to the difference test. Hence, it follows that $\Omega_{0,\mathsf{O}} \subset \Omega_{0,\mathsf{D}}$ if $\mu_{\mathsf{D}} = -2\ln(\mu_{\mathsf{O}} - 1)$, and the fail rate and success rate of the DTIA estimator will be higher than with OIA estimation.

If the precision as described by the vc-matrix $Q_{\hat{a}}$ is high the squared norms R_i, $i > 1$ will be large and the last term on the right-hand side of equation (5.63) will become small.

This means that the higher the precision, the more identical the aperture pull-in regions of the difference test and the optimal test become. It follows from equation (5.66) that then

$$\mu_{\mathsf{O}} \approx \exp\{-\frac{1}{2}\mu_{\mathsf{D}}\} + 1 \quad \text{or} \quad \mu_{\mathsf{D}} \approx -2\ln(\mu_{\mathsf{O}} - 1) \tag{5.67}$$

In figures 5.26 and 5.27 the aperture parameter μ_{O} is shown as function of the fail rate for different examples. The black solid lines show the μ_{O} that follow from the simulations, the dashed grey line show the μ'_{O} as approximated with (5.67) as function of the DTIA fail rate. For example 10_01, the precision is high and the results in table 5.2 already showed that the DTIA estimator performs almost identical to the OIA estimator. It could therefore be expected that $\mu_{\mathsf{O}} \approx \exp\{-\frac{1}{2}\mu_{\mathsf{D}}\} + 1$, which is indeed shown in the figure. On the other hand, for the other examples the DTIA estimator performs worse than the OIA estimator. Because of the poorer precision the terms $R_i, i > 3$, and thus the last term on the right-hand side of (5.63), cannot be ignored. Not surprisingly μ_{O} is clearly different from μ'_{O}.

In the bottom panels of the figures, the OIA success rates as function of the 'fixed' fail rate are shown in the case μ_{O} would be used, and in the case μ'_{O} would be used. So, in the latter case the success rate is shown as function of the DTIA fail rate which was used to determine μ'_{O}. As expected, only for example 10_01 approximately the same results are obtained.

From equation (5.66) follows that it is not possible to rewrite the OIA estimator in the same form as the ratio test (3.78), but both tests depend on R_1 and R_2, which is the reason that similar results can be obtained with RTIA and OIA estimation if it would be possible to choose the critical value of the ratio test in an appropriate way. In practice μ_{R} is often chosen equal to a fixed value and the ratio test should only be applied if the precision is high. Figure 5.24 shows on the left the aperture parameter μ_{R} as function of the fail rate for the different examples. It can be seen that for the examples where $n = 10$ a small fail rate will be obtained when $\mu_{\mathsf{R}} = 0.3$; in the case of poorer precisions on the other hand, it is not clear how the appropriate μ_{R} can be chosen such that the fail rate will be low enough, but at the same time such that the test is not too conservative.

The F-ratio test

$$\frac{\|\hat{e}\|^2_{Q_y} + R_1}{\|\hat{e}\|^2_{Q_y} + R_2} \leq \mu_{\mathsf{R}}$$

also looks at R_1 and R_2, but as explained this IA estimator is not really comparable with all others because it also depends on the float residuals $\hat{e}$. In practice the aperture parameter is often chosen equal to $\frac{1}{2}$ or as $1/F_\alpha(m-p, m-p, 0)$, see section 3.5. The panels on the right of figure 5.24 show the aperture parameter as function of the fail rate and as function of α in the case the F-distribution would be used to determine the aperture parameter. As explained, the latter choice is actually not correct. It follows that it is not obvious how α must be chosen, which confirms that it is better not to incorrectly use the F-distribution. With $\mu_{\mathsf{R}} = \frac{1}{2}$ the fail rate will generally be quite low.

The results of this section give an explanation for the fact that it is difficult to find a good criterion to choose the critical value for the ratio test and the difference test:

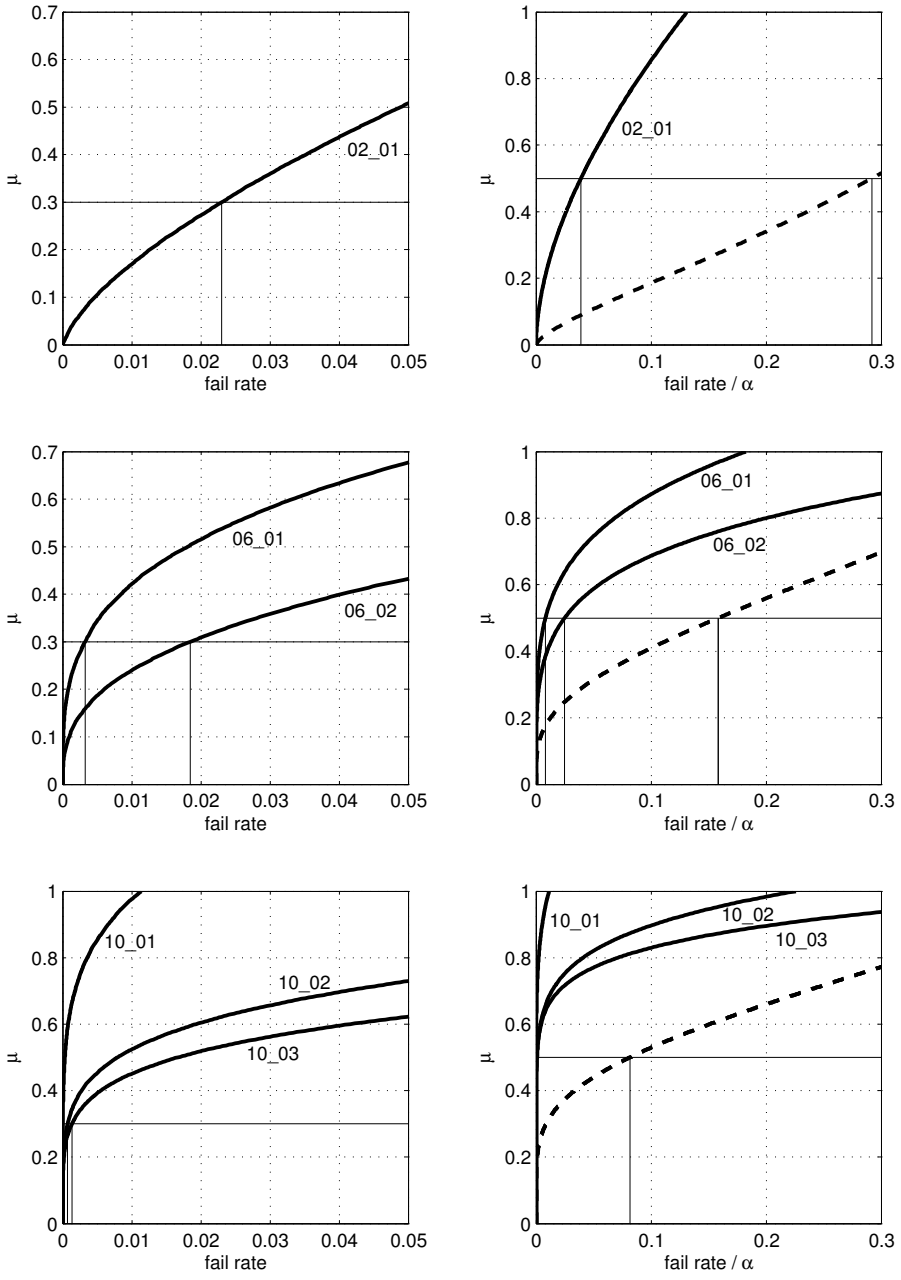

Figure 5.24: Aperture parameter μ as function of the fail rate. Left: RTIA; Right: F-RTIA. Top: example 02_01; Center: examples 06_xx; Bottom: examples 10_xx.

in order to guarantee a low fail rate the critical value should somehow depend on the distance to all integer vectors in the neighborhood of $\hat{a}$, where the distance is measured in the metric of $Q_{\hat{a}}$.

2-D examples

By comparing 2-D aperture pull-in regions of the RTIA, DTIA and OIA estimators, it is possible to get a better idea of the differences in performance of the three estimators.

Therefore three aperture pull-in regions for are plotted for different vc-matrices in figure 5.25. The following vc-matrices are considered:

$$\frac{2}{3}Q_{\hat{z}\_02\_01}, \quad Q_{\hat{z}\_02\_01}, \quad 4Q_{\hat{z}\_02\_01}, \quad Q_{\hat{z}\_02\_01} + \begin{bmatrix} 0.8 & 0 \\ 0 & 0 \end{bmatrix}$$

The regions correspond to the following three situations:

a : μ chosen such that for each estimator the fail rate is obtained for which the percentage of RTIA / DTIA solutions identical to the OIA solution, P_{id}, is minimum;
b : Ratio Test with fixed critical value of $\mu = 0.5$,
μ for DTIA and OIA chosen such that fail rates of all estimators are equal
c : all estimators with fixed fail rate of $P_f = 0.025$

It follows that the first situation, smallest P_{id} for the same fail rate, in general occurs when the OIA pull-in region just touches the ILS pull-in region, and thus the aperture parameter is large. The reason is that for a large aperture parameter the shape of the RTIA and DTIA pull-in regions start to resemble the shape of the ILS pull-in region and is larger in the direction of the corner points of the ILS pull-in regions where little probability mass is located. The OIA pull-in region, on the other hand, touches the ILS pull-in regions there where more probability mass is located in S_a, since the PDF of $\hat{a}$ is elliptically shaped. The smallest P_{id} is not obtained for a large aperture pull-in region when the precision is so high that the PDF of $\hat{a}$ is peaked since then there is little probability mass close to the boundaries of the ILS pull-in region.

If the aperture pull-in regions are chosen even larger than shown in figure 5.25, the OIA pull-in region is cut off by the ILS pull-in region. So, if μ is increased the OIA pull-in region cannot expand in all directions, and therefore the fail rate will not change much.

The aperture pull-in regions of the RTIA and DTIA estimators are especially different from those of the OIA estimator in the direction of the vertices of the ILS pull-in region. That is because in those directions there are two integers, $\check{a}_2$ and $\check{a}_3$ with the same distance to $\hat{a}$, but only $\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}$ is considered by the ratio test and difference test, whereas with OIA estimation the likelihood of all integers is considered.

Using a fixed critical value for the ratio test (b) results in very different and sometimes large fail rates, depending on the vc-matrix. The reason is that the size of the aperture pull-in regions is not adjusted if the precision changes. This can be seen for the first

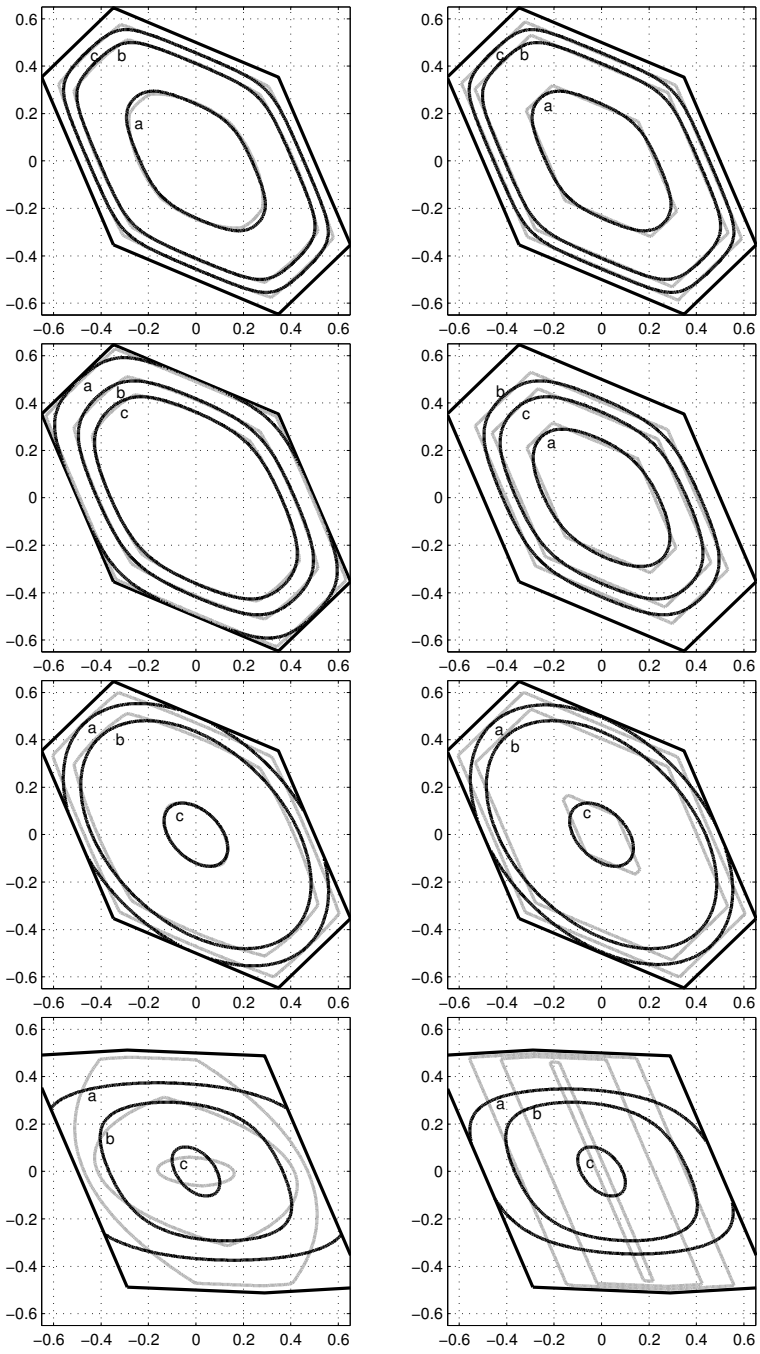

Figure 5.25: Aperture pull-in regions for: a) the μ that results in the smallest percentage of solutions identical to the optimal test; b) $\mu_R = 0.5$; c) $P_f = 0.025$. Left: OIA (black) and RTIA (grey); Right: OIA (black) and DTIA (grey). Top to bottom: $\frac{2}{3}Q_{\hat{z}\_02\_01}$; $Q_{\hat{z}\_02\_01}$; $4Q_{\hat{z}\_02\_01}$; $Q_{\hat{z}\_02\_01} + diag(0.8, 0)$.

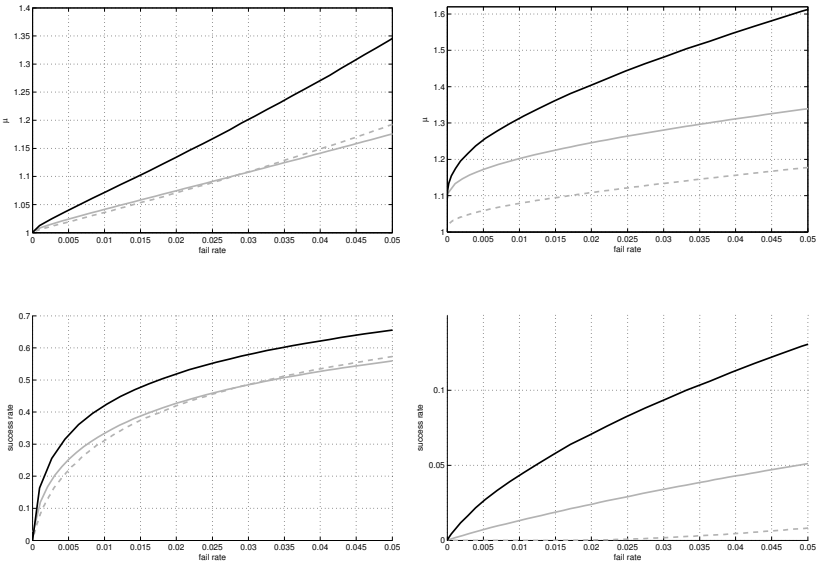

Figure 5.26: Aperture parameter μ_O (top) and success rate (bottom) as function of the fixed fail rate that is used to determine the aperture parameter. Black: μ_O determined from simulations; Dashed grey: $\mu'_O = \exp\{-\frac{1}{2}\mu_D\} + 1$; Solid grey: μ'_O approximated with IAB estimation. Left: example 06_01. Right: example 06_02.

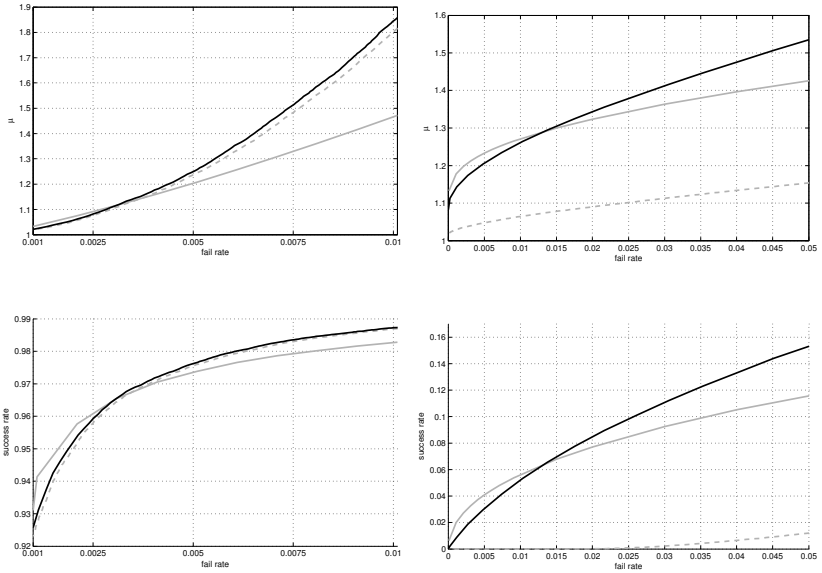

Figure 5.27: Aperture parameter μ_O (top) and success rate (bottom) as function of the fixed fail rate that is used to determine the aperture parameter. Black: μ_O determined from simulations; Dashed grey: $\mu'_O = \exp\{-\frac{1}{2}\mu_D\} + 1$; Solid grey: μ'_O approximated with IAB estimation. Left: example 10_01. Right: example 10_03.

three vc-matrices which are scaled versions of each other: the RTIA aperture pull-in regions corresponding to situation b are all equal.

The fixed fail rate approach works much better than using a fixed critical value for the ratio test, because the size of the aperture pull-in region is nicely adjusted in case of a fixed fail rate, as can be seen by comparing the results corresponding to situation c for the first three vc-matrices. In practice the choice of a fixed μ often works satisfactorily, but that is because either the value is chosen very conservative, cf. (Han 1997), or it is required that the ILS success rate is close to one before an attempt to fix the ambiguities is made.

The aperture pull-in regions of the DTIA and OIA estimators become more and more identical the higher the precision is, so that the relations in equation (5.67) become valid, as can be seen by comparing these regions for the first three vc-matrices.

5.8.3 IAB and IALS estimation

From table 5.2 it follows that IALS estimation may perform close to optimal, but IAB estimation does not perform so good. This can also be seen in the bottom panels of figures 5.28 and 5.29. For a certain fail rate, the success rates with IAB estimation are clearly lower than with IALS estimation.

The advantage of IAB estimation, however, is that exact evaluation of the fail rate and success rate is possible. Moreover, the results in table 5.2 indicate that for a fixed fail rate, the aperture parameters of both estimators do not differ that much. This can also be seen in the top panels of figures 5.28 and 5.29. For all examples $\mu_{\mathsf{B}} < \mu_{\mathsf{LS}}$. This gives rise to the idea to choose a maximum allowed fail rate, i.e. $P_f < \beta$, and compute the aperture parameter μ_{B} such that $P_{f,\mathsf{IAB}} = \beta$. Then apply IALS estimation with $\mu_{\mathsf{LS}} = \mu_{\mathsf{B}}$. For this choice it follows from equations (5.29) and (5.35) that the IALS success rate will always be larger than or equal to the IAB success rate. Based on the examples it can be expected that the IALS fail rate will then be smaller than β. However, this approach is not optimal since the corresponding success rate will also be lower than the one obtained with the IALS fail rate equal to β. This is illustrated with two examples in table 5.3. Moreover, it is not guaranteed that indeed $P_{f,\mathsf{IALS}}(\mu) \leq P_{f,\mathsf{IAB}}(\mu)$.

In section 5.7.1 an approximation of the OIA aperture parameter was given which makes use of the IAB aperture parameter. For examples 06_01, 06_02, 10_01 and 10_03 the results are shown in figures 5.26 and 5.27. The approximation does not work very well, and the results are often too conservative: the aperture pull-in regions are smaller than needed. Only for low fail rates the approximated aperture parameter is sometimes too large, so that the actual OIA fail rate will be larger than allowed.

5.9 Performance of IA estimation

The examples in the preceding section aimed at showing the potentials of IA estimation in general and comparing the different IA estimators. Based on the results it can be

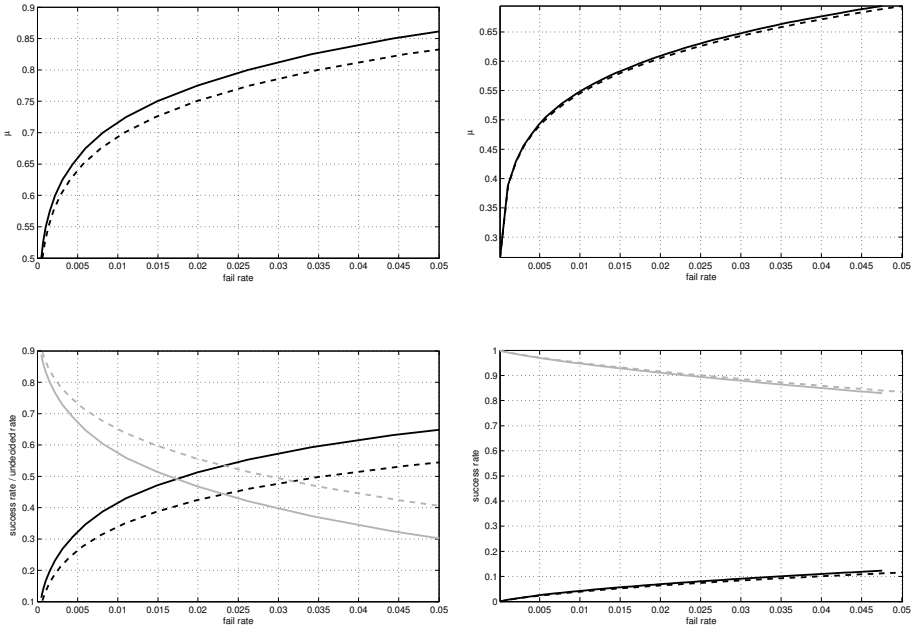

Figure 5.28: Comparison of IAB and IALS estimation, examples 06_01 (left) and 06_02 (right). Top: Aperture parameter μ as function of the fail rate for IAB (dashed) and IALS (solid) estimation. Bottom: Success rate (black) and undecided rate (grey) as function of the fail rate.

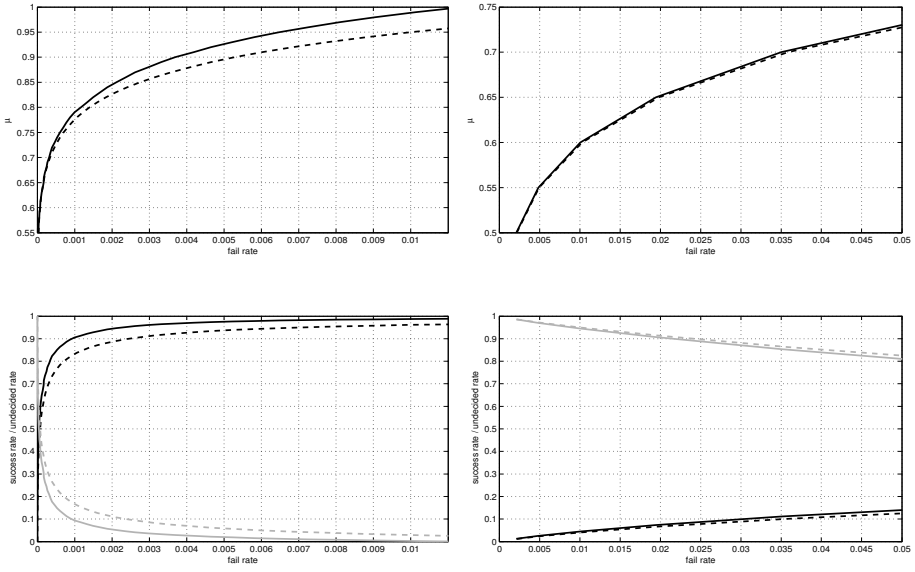

Figure 5.29: Comparison of IAB and IALS estimation, examples 10_01 (left) and 10_03 (right). Top: Aperture parameter μ as function of the fail rate for IAB (dashed) and IALS (solid) estimation. Bottom: Success rate (black) and undecided rate (grey) as function of the fail rate.

Table 5.3: Comparison of IAB and IALS estimation.

		IAB	IALS	
06_01	μ	0.5322	0.5322	0.5500
	P_f	0.0010	0.0008	0.0010
	P_s	0.1336	0.1458	0.1657
10_01	μ	0.7867	0.7867	0.8000
	P_f	0.0012	0.0010	0.0012
	P_s	0.8456	0.9023	0.9137

concluded that the DTIA, RTIA, and the IALS estimators perform almost as well as the OIA estimator. With the implementation as described in section 5.7.3 the DTIA and RTIA estimators have the advantage to be computationally simple and less demanding than the other two estimators. OIA estimation requires the approximation of $f_{\check{\epsilon}}(x)$ which may involve the determination of and summation over a very large integer set, see section 3.3. IALS estimation involves applying ILS estimation twice for each generated sample.

In this section, examples are worked out which illustrate the benefits of IA estimation with a fixed fail rate for practical use. Besides the OIA estimator only the RTIA estimator will be considered, because of its attractive properties described above, and because it is well-known and already often used in practice, albeit based on other acceptance criteria.

The vc-matrices of examples 06_02 and 10_03 are used. These vc-matrices correspond to a single epoch model. For the following epochs it is simply assumed that

$$Q_k = \frac{1}{k} Q_1,$$

with k the epoch number. This relation is valid when the satellite geometry is not changed.

The aperture parameters, success rates and fail rates as function of the number of epochs are determined using simulations. The following approaches are considered:

- Optimal IA estimation, fixed fail rate $P_f = 0.005$;
- Ratio Test IA estimation, fixed fail rate $P_f = 0.005$;
- Ratio Test, fixed critical value $\mu = \frac{1}{3}$;
- Ratio Test, fixed critical value $\mu = \frac{1}{2}$ **and**
  fixed solution is only accepted if $P_{s,\text{LS}} \geq 0.99$;
- fixed solution is only accepted if $P_{s,\text{LS}} \geq 0.995$.

If the ILS fail rate is smaller than 0.005, all solutions are accepted with the first two IA estimators, i.e. $\Omega_0 = S_0$. Note that the last approach also guarantees that the fail rate will never exceed 0.005.

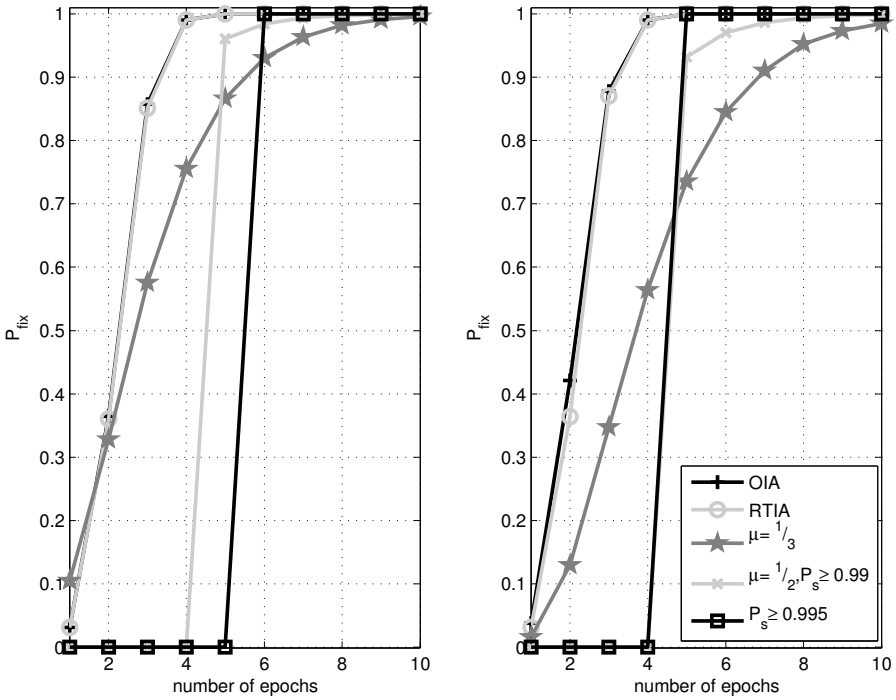

Figure 5.30: Probabilities that solution will be fixed as function of the number of epochs with 5 different tests. Left: 06_02, Right: 10_03.

For (near) real-time applications it is important that the time to first fix is as short as possible. The probability that a fix is made at a certain epoch equals:

$$P_{fix} = P_s + P_f \tag{5.68}$$

Figures 5.30 and 5.31 show P_{fix} and $P_{s|\bar{a}=z}$ obtained with the 5 approaches as function of the number of epochs. Obviously, IA estimation with a fixed fail rate result in the highest P_{fix}. Again, it follows that only with the fixed fail rate approach the performance of the RTIA estimator is close to optimal.

Applying the ratio test with a fixed critical value of $\frac{1}{3}$ (stars) may give comparable probabilities P_{fix} in the first epochs. Note that if the success rate is higher than the one obtained with the fixed fail rate, this implies that the corresponding fail rate is higher than 0.005. In later epochs the fixed critical value is far too conservative, so that the time to first fix may be unnecessarily long.

For the first epoch the probability $P_{s|\bar{a}=z}$ is always low for the examples considered here. The reason is that the success rate P_s is very low, and thus the approximation (5.8) is not valid. For later epochs $P_{s|\bar{a}=z}$ becomes reasonably high with all approaches, and is highest for the fixed critical value approach. Note, however, that this is only the case because this approach is too conservative; P_{fix} is very low in the first epochs, so that if the solution is fixed it can be expected to be correct since the aperture space, and thus

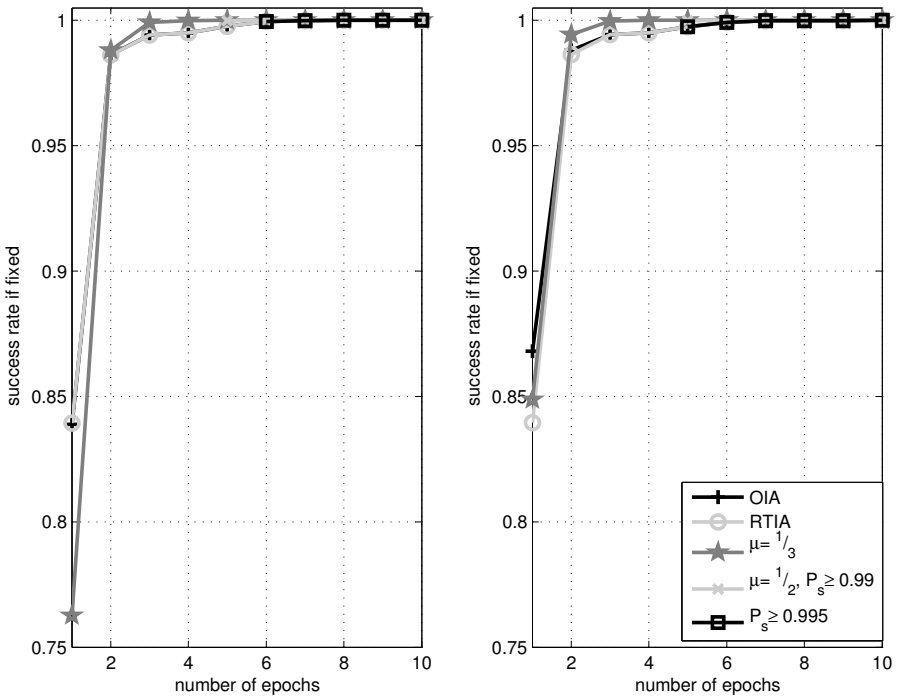

Figure 5.31: Success rates conditioned on $\bar{a} = z$ as function of the number of epochs with 5 different tests. Left: 06_02, Right: 10_03.

also the fail rate, is very small.

With the two approaches where it is required that the ILS success rate exceeds a certain limit (x-signs and squares) no attempt is made to fix the solution in the first epochs. As soon as a fix is made, the probability $P_{s|\bar{a}=z}$ is high.

With the OIA and the RTIA estimator with a fixed fail rate there is a very high probability that the solution will be fixed much faster than with the other approaches, and the RTIA estimator performs close to optimal. At the same time it is guaranteed that the fail rate will not exceed the user-defined threshold. This is not the case if the ratio test with a fixed critical value is used. It can be concluded that the performance of the ratio test with a fixed fail rate is much better than with a fixed critical value, especially when more epochs of data become available. The reason is that using a fixed critical value implies that the aperture pull-in region Ω_0 does not change as function of the number of epochs. Recall that the success rate is given by:

$$P_s = \int_{\Omega_0} f_{\hat{a}}(x+a)dx$$

So, if more epochs of data become available, the success rate only increases because $f_{\hat{a}}(x+a)$ becomes more peaked due to the improved precision. Using a fixed fail rate, on the other hand, means that Ω_0 becomes larger when the precision improves. This

implies that the success rate increases more than with the fixed critical value approach, since $f_{\hat{a}}(x+a)$ is the same, but the integration region is larger.

5.10 Summary

In practice validation of the integer ambiguity solution is taken care of by using test statistics, which lack a sound theoretical basis. The problem is then the choice of the critical value.

As an alternative, in this chapter the concept of integer aperture estimation with a fixed fail rate was presented. The advantages are:

- It is an overall approach: estimation and validation are not considered as separate problems.
- An exact and overall probabilistic evaluation of the final solution is possible.
- The user only needs to choose the maximum allowable fail rate, which is exactly the parameter on which a user wants to put restrictions. Moreover, this means that the size of the acceptance region is really based on the model at hand. Choosing a fixed critical value or a critical value based on incorrect assumptions on the distribution of the test statistic, as is common practice, will only work when the precision is high, and may even then result in a test that is either too conservative or too optimistic.
- The discrimination tests – ratio tests, difference test, projector test – as used in practice belong to the class of integer aperture estimators, and the performance of the ratio test and difference test is shown to be close to optimal. So, based on the theory of Integer Aperture inference the theoretical foundation as well as the practical relevance of the ratio test and difference test have been shown.
- Using the fixed fail rate approach implies that the critical value depends on the model (and thus the precision) at hand. This means that the time to first fix will be shorter, and at the same time it is guaranteed that the probability of incorrect fixing is below a user-defined threshold.

Several new IA estimators have been presented; the optimal IA estimator is one of them. This estimator maximizes the success probability for a given fail rate. An overview of all IA estimators is given in table 5.4. Unfortunately, for most IA estimators it is difficult to determine the aperture parameter.

Simulations were used to evaluate and compare the IA estimators. It follows that the ratio test IA estimator (RTIA), the difference test IA estimator (DTIA), and the IA least-squares estimator (IALS) have a close to optimal performance. In practice often the RTIA estimator is used, because it is known to work well if the success rate is high. From the comparison with the OIA estimator, it is now explained why this estimator and the DTIA estimator perform so well. The close to optimal performance of the IALS

Table 5.4: Overview of IA estimators. The last column gives the methods that can be used for the determination of the aperture parameter μ in the case of fixed fail rate IA estimation.

	definition aperture space:		determination μ
EIA	ellipsoid	$\parallel \hat{a} - \check{a} \parallel^2_{Q_{\hat{a}}} \leq \mu^2$	root finding method
RTIA	ratio test	$\frac{\|\hat{a}-\check{a}\|^2_{Q_{\hat{a}}}}{\|\hat{a}-\check{a}_2\|^2_{Q_{\hat{a}}}} \leq \mu$	empirical/simulation
F-RTIA	F-ratio test	$\frac{\|\hat{e}\|^2_{Q_y}+\|\hat{a}-\check{a}\|^2_{Q_{\hat{a}}}}{\|\hat{e}\|^2_{Q_y}+\|\hat{a}-\check{a}_2\|^2_{Q_{\hat{a}}}} \leq \mu$	empirical/simulation
DTIA	difference test	$\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}} - \|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}} \geq \mu$	empirical/simulation
PTIA	projector test	$\frac{(\check{a}_2-\check{a})^T Q^{-1}_{\hat{a}}(\hat{a}-\check{a})}{\|\check{a}_2-\check{a}\|_{Q_{\hat{a}}}} \leq \mu$	empirical/simulation
IAB	bootstrapping	$\frac{1}{\mu}\check{\epsilon}_{\mathsf{B}} \in S_{0,\mathsf{B}}$	root finding method
IALS	least-squares	$\frac{1}{\mu}\check{\epsilon}_{\mathsf{LS}} \in S_{0,\mathsf{LS}}$	$= \mu_{\mathsf{B}}$/simulation
OIA	optimal	$\frac{f_{\check{\epsilon}}(\check{\epsilon}_{\mathsf{LS}})}{f_{\hat{a}}(\check{\epsilon}_{\mathsf{LS}}+a)} \leq \mu$	simulation

estimator is also not surprising, since it makes use of the scaled ILS pull-in region as aperture space, and the ILS estimator is known to be optimal in the class of admissible integer estimators.

Which IA estimator to use?

The RTIA and DTIA estimator are easy to apply and computationally simple, since the squared norms of the ambiguity residuals are already available after computing the ILS solution. However, these estimators are only close to optimal if an appropriate aperture parameter is used based on a fixed fail rate; and that is only possible by using simulations.

The IALS estimator might be a good alternative: close to optimal performance, and easy to apply. The aperture parameter could be chosen equal to that of the IAB estimator for a given fail rate. Unfortunately, this will make the IALS estimator less optimal. If simulations are used for the determination of the aperture parameter, the computational efficiency of the IALS estimator will be much less than that of the RTIA and DTIA estimators, since ILS estimation has to be applied twice for each sample.

In theory, the OIA estimator is of course to be preferred since it maximizes the success rate. However, computationally it is the most complex estimator because the probability density of the ambiguity residual must be computed. Furthermore, it is only possible to determine the aperture parameter for a given fail rate using simulations. If the aperture parameter is approximated using the IAB estimator, the resulting estimator will not be optimal.

6 Conclusions and recommendations

6.1 Integer estimation and validation

The goal of this thesis was to evaluate different approaches to the integer estimation and validation problem. In this thesis three approaches were considered, based on three classes of integer estimators. The relationships between the estimators are shown in the Venn diagram of figure 6.1.

The smallest, most restrictive class is the class of admissible integer estimators. Integer rounding, bootstrapping, and integer least-squares are the well-known admissible estimators. From this class, integer least-squares is the optimal estimator, since it maximizes the probability of correct integer estimation. Therefore, almost all ambiguity resolution methods currently used in practice are based on this estimator. A parameter resolution theory can, however, not be considered complete without the appropriate measures to validate the solution. For that purpose, many integer validation tests have been proposed in literature, but all of these tests lack a sound theoretical foundation and most tests are based on the invalid assumption that the integer ambiguity candidates are deterministic.

The largest class of integer estimators is the class of best integer equivariant (BIE) estimators. This BIE estimator is in some sense always superior to its float and fixed counterparts: it is best in the sense that it minimizes the mean squared errors of the estimators. This is a weaker performance criterion than that of the integer least-squares estimator, the maximization of the success rate. This might be an advantage if the success rate is not close to one, since then the fixed ambiguities may not be considered deterministic and the integer validation procedures that are currently available should not be used. So, in that case the ambiguities cannot be fixed reliably.

The BIE ambiguity estimator is equal to a weighted sum over all integer vectors, and the weights depend on the probability density function of the float ambiguities. Therefore, the BIE estimator can be considered as a compromise between the float and fixed solution. These are approximated in the extreme cases of very bad or very good precision respectively. The disadvantages of BIE estimation are that it does still not provide an exact probabilistic evaluation of the final solution, and that it is computationally complex. Moreover, the results in this chapter indicate that only in a limited number of cases the BIE estimator performs obviously better than either the float or the fixed estimator.

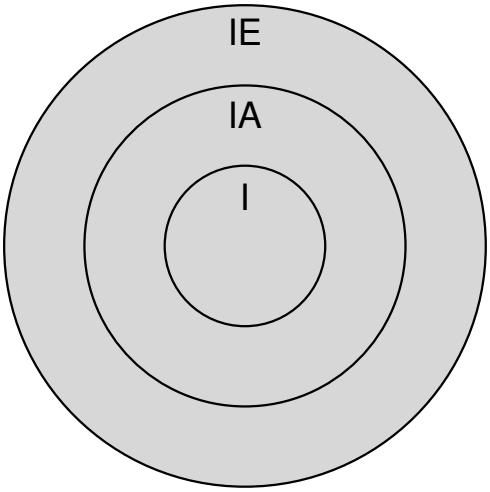

Figure 6.1: The set of relationships between integer estimators (I), integer aperture estimators (IA), and integer equivariant estimators (IE): I⊂IA⊂IE.

Admissible integer estimation and BIE estimation both cannot provide us with satisfactory measures to evaluate the probabilistic properties of the final solution. Therefore, the third class of integer estimators was introduced: the class of integer aperture (IA) estimators.

IA estimation is an overall approach of integer estimation and validation. This is the case because three a priori probabilities are distinguished: the success rate, the fail rate, and the undecided rate. Around each integer an identically shaped acceptance region, referred to as the aperture pull-in region, is defined, with the size determined by the choice of a maximum allowed fail rate. With respect to the shape of the regions different choices are possible. Some 'simple' choices are using ellipsoidal regions, or the down-scaled pull-in regions of integer bootstrapping or integer least-squares. But other choices are possible. The Optimal IA estimator is defined such that the success rate is maximized for a given fixed fail rate.

Advantages of IA estimation with a fixed fail rate are:

- It is an overall approach: estimation and validation are not considered as separate problems.
- An exact and overall probabilistic evaluation of the final solution is possible.
- The user only needs to choose the maximum allowable fail rate, which is exactly the parameter on which a user wants to put restrictions. Moreover, this means that the size of the acceptance region is really based on the model at hand. Choosing a fixed critical value or a critical value based on incorrect assumptions on the distribution of the test statistic, as is common practice, will only work when the precision is high, and may even then result in a test that is either too conservative or too optimistic.

- The discrimination tests – ratio tests, difference test, projector test – as used in practice belong to the class of integer aperture estimators, and the performance of the ratio test and difference test is shown to be close to optimal. So, based on the theory of Integer Aperture inference the theoretical foundation as well as the practical relevance of the ratio test and difference test have been shown.
- Using the fixed fail rate approach implies that the critical value depends on the model (and thus the precision) at hand. This means that the time to first fix will be shorter, and at the same time it is guaranteed that the probability of incorrect fixing is below a user-defined threshold.

A problem with IA estimation is the determination of the aperture parameter for a given fail rate. For that purpose simulations must be used. More research is needed on the consequences on the computation times and on the performance.

In theory, the Optimal IA estimator is of course to be preferred since it maximizes the success rate. However, computationally it is the most complex estimator because the probability density of the ambiguity residual must be computed.

The ratio test and difference test IA estimator are easier to apply and computationally simple. These estimators have been shown to perform close to optimal, if and only if the appropriate aperture parameter is determined by the fixed fail rate. The ratio test is already often used in practice and seems therefore a logical choice.

The IALS estimator might be a good alternative: close to optimal performance, and easy to apply, although computationally much more demanding than the ratio test and difference test IA estimators.

More research is needed on the implementation aspects and on how the IA method will work in practice for different applications. With respect to the implementation aspects it will be important to find an efficient way of determining the aperture parameter such that close to optimal solutions will be obtained. The approach followed in this thesis was to use simulations with a fairly small number of samples. It must be investigated how this will work in practice. Alternative methods would be to somehow determine empirical values of the aperture parameter that can be used under certain conditions, e.g. depending on the ILS success rate or the number of ambiguities. Or it may be possible to obtain a good approximation of the appropriate aperture parameter based on information from previous epochs.

6.2 Quality of the baseline estimators

Integer ambiguity validation focuses on the question whether or not the ambiguities should be fixed. This decision is based on the residuals of the float and fixed solution. A user, however, will be mainly interested in the impact of either decision on the baseline solution. Therefore, in this thesis several baseline probabilities were evaluated using simulations.

The first probability is the probability that a baseline estimate will be within a certain

convex region symmetric with respect to the true baseline. It followed that the probability that the fixed solution is close to the true solution is always larger than for any of the other baseline estimators. At the same time, the fixed solution also has the highest probability of being far from the true solution. The opposite is true for the float estimator: smallest probability of being close to the true solution, but also small probability that it will be far off. The probabilities of both the BIE and the IA estimators generally are inbetween those of the fixed and float estimators. So, as could be expected the BIE and IA estimators take the best of the float and fixed estimators.

Another probability considered here is the probability that a baseline estimator is better than the corresponding float estimator. Especially if IA estimation with a small fixed fail rate is used, this probability will be close to one and larger than for any of the other estimators.

A problem is that the probabilities cannot be evaluated exactly, and no strict lower and upper bounds are available, so that simulations are needed. It would be interesting to have an easy-to-evaluate measure of the quality of the baseline solution. This is a point of further research.

6.3 Reliability of the results

In this thesis simulations were used in order to assess the theoretical performance of different estimators and to make comparison possible, since one can be sure that the simulated data correspond to the mathematical model that is used. In practice, one cannot be sure that this will be the case, whereas the reliability of the ambiguity resolution results depends on the validity of the input. Only if one can assume that the input is correct, one can trust the outcome of integer estimation and validation.

The input of the ambiguity resolution process consists of the float ambiguity estimates and their vc-matrix, and thus depends on the functional as well as the stochastic model that is used. Furthermore, it is important that the quality of the float solution is evaluated, such that errors due to for example multipath and cycle slips are eliminated.

In practice the stochastic model is often assumed to be correct, although recent studies have shown that in many cases a refinement is possible. Possible refinements are the application of satellite elevation dependent weighting of the observations, and considering cross-correlation of the different observations. This means an appropriate stochastic model must be determined for the receiver at hand.

Another assumption that was made is that the variance factor of unit weight, σ^2, is known. If, on the other hand, it is assumed that this variance factor is not known, it must be estimated a posteriori by means of variance component estimation. It must be investigated how this affects the results presented in this thesis.

Furthermore, it is assumed that the observations are normally distributed. The results in this thesis are all based on this assumption, although it would also be possible to apply the theory to parameters with different distributions. This is also an issue open for further research.

6.4 Bias robustness

In the preceding section it was stressed that it is important to eliminate errors in the float solution as much as possible. In that respect it would be interesting to investigate how 'common' biases, such as multipath, outliers in the data, and cycle slips, propagate into the final solution if they are not corrected for or eliminated. It is namely possible that due to the bias the ambiguity is estimated to the wrong integer with IA estimation, which might result in a baseline estimate that has a larger bias than the float baseline estimate.

A Mathematics and statistics

A.1 Kronecker product

The Kronecker product is defined as:

$$M \otimes N = \begin{pmatrix} m_{11}N & \cdots & m_{1q}N \\ \vdots & & \vdots \\ m_{p1}N & \cdots & m_{pq}N \end{pmatrix} \qquad \text{with } M = [m_{ij}] \tag{A.1}$$

It has the following properties:

$$\begin{aligned} (M \otimes N)^T &= M^T \otimes N^T \\ (M \otimes N)^{-1} &= M^{-1} \otimes N^{-1} \\ (M_1 \otimes N_1)(M_2 \otimes N_2) &= (M_1 M_2 \otimes N_1 N_2) \end{aligned}$$

A.2 Parameter distributions

A.2.1 The normal distribution

The probability density function (PDF), $f_{\underline{x}}(x)$, of a normally distributed $n \times 1$ random vector $\underline{x}$ is given by:

$$f_{\underline{x}}(x) = (2\pi)^{-\frac{n}{2}} |Q|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(x-\mu)^T Q^{-1}(x-\mu)\} \tag{A.2}$$

with Q a $n \times n$ positive definite matrix, and μ an $n \times 1$ vector. Once Q and μ are given, the normal distribution is thus completely specified. The following notation is used:

$$\underline{x} \sim N(\mu, Q) \tag{A.3}$$

The expectation, $E\{\underline{x}\}$, and the dispersion (or variance-covariance matrix), $D\{\underline{x}\}$, of $\underline{x}$ are:

$$E\{\underline{x}\} = \mu \quad \text{and} \quad D\{\underline{x}\} = Q \tag{A.4}$$

Let $\underline{x}$ be normally distributed with expectation and dispersion as in (A.4), and let $\underline{y}$ be defined as the $m \times 1$ random vector given by $\underline{y} = A\underline{x} + a$. Then according to the propagation law of the mean and of variances:

$$E\{\underline{y}\} = A\mu + a \quad \text{and} \quad D\{\underline{x}\} = AQA^T \tag{A.5}$$

and thus:

$$\underline{y} \sim N(A\mu + a, AQA^T) \tag{A.6}$$

The cumulative normal distribution is given by:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}v^2\}dv \tag{A.7}$$

Some useful remarks:

$$\int_{x}^{y} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}v^2\}dv = \Phi(y) - \Phi(x) \tag{A.8}$$

$$\Phi(-x) = 1 - \Phi(x) \tag{A.9}$$

A.2.2 The χ^2-distribution

A scalar random variable, $\underline{x}$, has a non-central χ^2-distribution with n degrees of freedom and non-centrality parameter λ, if its PDF is given as:

$$f_{\underline{x}}(x) = \begin{cases} \exp\{-\frac{\lambda}{2}\} \sum_{j=0}^{\infty} \frac{(\frac{\lambda}{2})^j x^{\frac{n}{2}+j-1} \exp\{-\frac{x}{2}\}}{j! 2^{\frac{n}{2}+j} \Gamma(\frac{n}{2}+j)} & \text{for } 0 < x < \infty \\ 0 & \text{for } x \leq 0 \end{cases} \tag{A.10}$$

with the gamma function:

$$\Gamma(x) = \int_{0}^{\infty} t^{x-1} \exp\{-t\}dt, \quad x > 0 \tag{A.11}$$

The values of $\Gamma(x)$ can be determined using $\Gamma(x+1) = x\Gamma(x)$, with $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ and $\Gamma(1) = 1$.

The following notation is used:

$$\underline{x} \sim \chi^2(n, \lambda) \tag{A.12}$$

If $\lambda = 0$, the distribution is referred to as the central χ^2-distribution.

The expectation, $E\{\underline{x}\}$, and the dispersion, $D\{\underline{x}\}$, of $\underline{x} \sim \chi^2(n,\lambda)$ are:

$$E\{\underline{x}\} = n + \lambda \quad \text{and} \quad D\{\underline{x}\} = 2n + 4\lambda \tag{A.13}$$

If $\underline{x} \sim N(\mu, Q)$ and $\underline{y} = \underline{x}^T Q^{-1} \underline{x}$ then:

$$\underline{y} \sim \chi^2(n,\lambda) \quad \text{with} \quad \lambda = \mu^T Q^{-1} \mu \tag{A.14}$$

A.2.3 The F-distribution

A scalar random variable, $\underline{x}$, has a non-central F-distribution with m and n degrees of freedom and non-centrality parameter λ, if its PDF is given as:

$$f_{\underline{x}}(x) = \begin{cases} \exp\{-\frac{\lambda}{2}\} \sum_{j=0}^{\infty} \frac{(\frac{\lambda}{2})^j x^{\frac{m}{2}+j-1} m^{\frac{m}{2}+j} n^{\frac{n}{2}} \Gamma(\frac{m}{2}+\frac{n}{2}+j)}{j!\Gamma(\frac{m}{2}+j)\Gamma(\frac{n}{2})(n+mx)^{\frac{m}{2}+\frac{n}{2}+j}} & \text{for } 0 < x < \infty \\ 0 & \text{for } x \leq 0 \end{cases} \tag{A.15}$$

The following notation is used:

$$\underline{x} \sim F(m,n,\lambda) \tag{A.16}$$

If $\lambda = 0$, the distribution is referred to as the central F-distribution.

The expectation, $E\{\underline{x}\}$, and the dispersion, $D\{\underline{x}\}$, of $\underline{x} \sim F(m,n,\lambda)$ are:

$$E\{\underline{x}\} = \frac{n}{n-2} \ (n > 2) \quad \text{and} \quad D\{\underline{x}\} = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \ (n > 4) \tag{A.17}$$

If $\underset{m\times 1}{\underline{u}} \sim N(u, Q_u)$, $\underset{n\times 1}{\underline{v}} \sim N(v, Q_v)$, and $\underline{u}$ and $\underline{v}$ are uncorrelated, then

$$\underline{x} = \frac{\underline{u}^T Q_u^{-1} \underline{u}/m}{\underline{v}^T Q_v^{-1} \underline{v}/n}$$

is distributed as:

$$\underline{x} \sim F(m,n,\lambda) \quad \text{with} \quad \lambda = u^T Q_u^{-1} u \tag{A.18}$$

The distribution of $\underline{x} = \underline{u}^T Q_u^{-1} \underline{u}/m$ is given as: $\underline{x} \sim F(m,\infty,\lambda)$.

A.2.4 Student's t-distribution

A scalar random variable, $\underline{x}$, has a Student's t-distribution with n degrees of freedom, if its PDF is given as:

$$f_{\underline{x}}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \quad \text{for } -\infty < x < \infty \tag{A.19}$$

The following notation is used:

$$\underline{x} \sim t(n) \tag{A.20}$$

If $\underline{u} \sim N(0,1)$, and $\underline{v} \sim \chi^2(n,0)$, then

$$\underline{x} = \frac{\underline{u}}{\sqrt{\underline{v}/n}} \sim t(n)$$

A.3 Numerical root finding methods

In order to find the root of a function $f(x)$, several methods can be used. The four methods considered here are the bisection method, the secant method, the false position method, and the Newton-Raphson method.

A.3.1 Bisection method

The bisection method is very simple and will never fail if the initial interval is chosen such that it contains a root. However, the method may work very slowly. The procedure is as follows.

Choose the interval $[x_0^-, x_0^+]$ such that $f(x_0^-)f(x_0^+) < 0$.

Choose the tolerance ε.

Compute the function value at the middle of the interval, i.e. choose $x_{k+1} = \frac{x_0^+ - x_0^-}{2}$, and evaluate whether or not:

$$|f(x_{k+1})| < \varepsilon, \quad k = 0, 1, 2, \ldots$$

If yes, the root is found, otherwise:

$$\begin{cases} \text{if } f(x_{k+1})f(x_k^+) < 0: & x_{k+1}^- = x_{k+1},\ x_{k+1}^+ = x_k^+ \\ \text{else} & x_{k+1}^- = x_k^-,\ x_{k+1}^+ = x_{k+1} \end{cases}$$

and continue.

A.3.2 Secant method

The secant method converges faster to its solution than the bisection method, but it may not converge at all if the function is not sufficiently smooth. The method namely assumes that the function is approximately linear in the chosen interval and uses the zero-crossing of the line that connects the limits of the interval as a new reference point. The procedure is as follows.

Choose the interval $[x_0, x_1]$ such that $f(x_0)f(x_1) < 0$.

Choose the tolerance ε.

Determine the zero-crossing of the line connecting the limits of the interval:

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k), \quad k = 1, 2, 3, \ldots$$

Compute the function value at x_{k+1}, and evaluate whether or not:

$$|f(x_{k+1})| < \varepsilon$$

If yes, the root is found, otherwise continue.

A.3.3 False position method

The false position method works similar to the secant method, but instead of the most recent estimates of the root, the false position method uses the most recent estimate and the next recent estimate which has an opposite sign in the function value. The procedure is then as follows.

Choose the interval $[x_0^-, x_0^+]$ such that $f(x_0^-) f(x_0^+) < 0$.

Choose the tolerance ε.

Determine the zero-crossing of the line connecting the limits of the interval:

$$x_{k+1} = x_k^+ - \frac{x_k^+ - x_k^-}{f(x_k^+) - f(x_k^-)} f(x_k^+), \quad k = 0, 1, 2, \ldots$$

Compute the function value at x_{k+1}, and evaluate whether or not:

$$|f(x_{k+1})| < \varepsilon$$

If yes, the root is found, otherwise:

$$\begin{cases} \text{if } f(x_{k+1}) f(x_k^+) < 0: & x_{k+1}^- = x_{k+1},\ x_{k+1}^+ = x_k^+ \\ \text{else} & x_{k+1}^- = x_k^-,\ x_{k+1}^+ = x_{k+1} \end{cases}$$

and continue.

A.3.4 Newton-Raphson method

The Newton-Raphson method determines the slope of the function at the current estimate of the root and uses the zero-crossing of the tangent line as the next reference point. The procedure is as follows.

Choose the initial estimate x_0.

Choose the tolerance ε.

Determine the zero-crossing of the tangent line of the function at x_k:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \ldots$$

Compute the function value at x_{k+1}, and evaluate whether or not:

$$|f(x_{k+1})| < \varepsilon$$

If yes, the root is found, otherwise continue.

The Newton-Raphson method is in general much more efficient than other 'simple' methods such as the bisection method, but it can only be used when it is possible to compute the derivative of $f(x)$ at the reference point. It is, however, possible to simplify the method by replacing $f'(x_k)$ by a constant, but this may corrupt the speed of convergence.

A.3.5 Matlab function `fzero`

In this research the Matlab® function `fzero` was used in order to find the root of a function. The algorithm that is used is a combination of bisection, the secant method and inverse quadratic interpolation, and is based on the Algol 60 version described in (Brent 1973).

B Simulation and examples

Throughout this work several examples are based on simulations for which the vc-matrix of the float solution is required as input. In this appendix the method of simulation and the underlying models for the input are described.

B.1 Simulation

Simulation is a useful method for evaluation of the success rate and fail rate. Although in practice it would be too time-consuming, the method is used in this research in order to evaluate the results of different ambiguity resolution methods. Simulation goes as follows. Since it is assumed that the float solution is normally distributed, the probabilities are independent of the mean, so one can use $N(0, Q)$ and draw samples from this distribution.

For example, if only float ambiguities need to be generated, the first step is to use a random generator to generate n independent samples from the univariate standard normal distribution $N(0, 1)$, and then collect these in a vector s. This vector is transformed by means of $\hat{a} = Gs$, with G equal to the Cholesky factor of $Q_{\hat{a}} = GG^T$. The result is a sample $\hat{a}$ from $N(0, Q_{\hat{a}})$, and this sample is used as input for the integer estimator that is considered. If the output of this estimator equals the null vector, then it is correct, otherwise it is incorrect. This process can be repeated N number of times, and one can count how many times the null vector is obtained as solution, say N_s times, and how often the outcome equals a nonzero *integer* vector, say N_f times. The approximations of the success rate and fail rate follow than as:

$$P_s = \frac{N_s}{N}, \quad P_f = \frac{N_f}{N}$$

In order to get good approximations, the number of samples N must be chosen sufficiently large, see (Teunissen 1998c).

The random generator used in this research is the Matlab® routine `mvnrnd` from the Statistics toolbox.

Table B.1: Time, location, number of satellites, ionospheric standard deviations and empirical ILS success rates for the different $Q_{\hat{a}\_nn\_xx}$.

nn_xx	σ_I [cm]	success rate	
02_01	0	0.869	$Q_{\hat{a}} = \begin{bmatrix} 4.9718 & 3.8733 \\ 3.8733 & 3.0188 \end{bmatrix}$, $Q_{\hat{z}} = \begin{bmatrix} 0.0865 & -0.0364 \\ -0.0364 & 0.0847 \end{bmatrix}$
06_01	0	0.818	10:05 am, 4° East, 52° North
06_02	1	0.442	4 visible satellites
08_01	0	0.999	10:15 am, 4° East, 52° North
08_02	1	0.874	5 visible satellites
10_01	1	0.989	17:00 am, 52° East, 4° North
10_02	2	0.774	6 visible satellites
10_03	3	0.477	

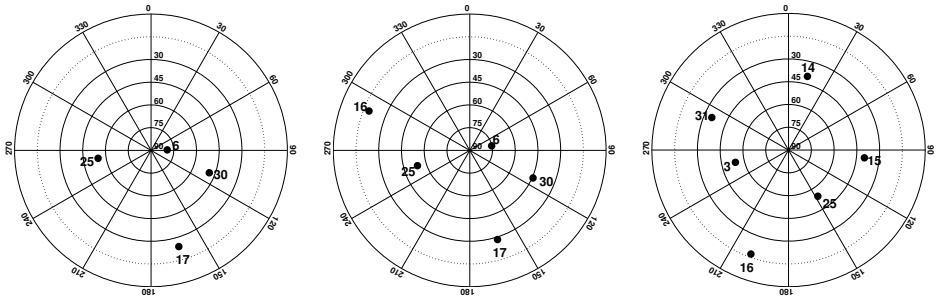

Figure B.1: Skyplots corresponding to examples 06_xx (left), 08_xx (center), and 10_xx (right).

B.2 Examples

The names assigned to the different vc-matrices of the float ambiguities all have the form $Q_{\hat{a}\_\mathbf{nn\_xx}}$ and $Q_{\hat{z}\_\mathbf{nn\_xx}}$. The number nn is equal to the number of ambiguities n, and xx is an identification number.

The two-dimensional example ($n = 2$) corresponds to the double difference geometry-free model for dual-frequency GPS for two satellites. The higher-dimensional examples are all based on the single epoch geometry-based model for dual-frequency GPS. These models are all set up based on the Yuma almanac for GPS week 184, 19-FEB-2003; a cutoff elevation of 15°; undiffenced code and phase standard deviations of 30 cm and 3 mm respectively.

C Theory of BIE estimation

In chapter 4 the theory of Best Integer Equivariant estimation was developed. In this appendix a more detailed description of the derivations is given based on (Teunissen 2003f).

C.1 Integer equivariant ambiguity estimation

Starting point is the integer estimator of the unknown integer ambiguities $a \in \mathbb{Z}^n$. Let this estimator be $S(\hat{a})$ with $S : \mathbb{R}^n \mapsto \mathbb{Z}^n$. The new class of integer estimators to be defined here will be larger than the class of admissible integer estimators of definition 3.1.1. Only the third condition of that definition will be considered, i.e. $S(\hat{a} - z) + z = S(\hat{a})$, $\forall z \in \mathbb{Z}^n$. This means that the result of integer estimation should not change when first an arbitrary integer z is removed from the float solution, then apply the integer estimator, and then restore the integer. So, the requirement is equivalent to:

$$S(x + z) = S(x) + z \quad \forall x \in \mathbb{R}^n, z \in \mathbb{Z}^n \tag{C.1}$$

Estimators that satisfy this property will be called integer equivariant (IE).

Lemma C.1.1

$$S(x + z) = S(x) + z \quad \forall x \in \mathbb{R}^n, z \in \mathbb{Z}^n \Longleftrightarrow S(x) = x + g(x) \tag{C.2}$$

with $g(x)$ periodic: $g(x + z) = g(x)$.

Proof:
If $S(x) = x + g(x)$ with $g(x + z) = g(x)$, then $S(x + z) = S(x) + z$.
If $S(x + z) = S(x) + z$, then $g(x) = x - g(x)$ is periodic and $S(x) = x + g(x)$.

□

C.2 Integer equivariant unbiased ambiguity estimation

It follows that:

$$a = E\{S(\hat{a})\} = E\{\hat{a} + g(\hat{a})\} = a + E\{g(\hat{a})\}$$

So that:

$$0 = E\{g(\hat{a})\} = \int g(x) f_{\hat{a}}(x) dx = \int g(x) f_{\hat{a}}(x) dx$$

Thus if the PDF $f_{\hat{a}}(x+a)$ is symmetric with respect to the origin, which is true in the case of normally distributed data, then $g(-x) = -g(x)$ implies unbiasedness.

C.3 Best integer equivariant unbiased ambiguity estimation

The least mean square estimator will be considered as the *best* integer equivariant estimator.

Theorem C.3.1

$$\hat{S}(x) = x - \frac{\sum\limits_{z \in \mathbb{Z}^n} (x+z-a) f_{\hat{a}}(x+z)}{\sum\limits_{z \in \mathbb{Z}^n} f_{\hat{a}}(x+z)} = \frac{\sum\limits_{z \in \mathbb{Z}^n} z f_{\hat{a}}(x+a-z)}{\sum\limits_{z \in \mathbb{Z}^n} f_{\hat{a}}(x+a-z)} \tag{C.3}$$

is the unique minimizer of $\int \|S(x)-a\|^2_{Q_{\hat{a}}} f_{\hat{a}}(x) dx$ within the class of integer equivariant estimators, $S(x+z) = S(x) + z$

Proof:
Let $\|\cdot\|^2 = (\cdot)^T Q_{\hat{a}}^{-1}(\cdot)$. Then:

$$\begin{aligned}
&\int \|S(x)-a\|^2 f_{\hat{a}}(x) dx \\
&\qquad = \int \|x-a+g(x)\|^2 f_{\hat{a}}(x) dx \\
&\qquad = \int \|x-a\|^2 f_{\hat{a}}(x) dx + \int \{2g(x)^T Q_{\hat{a}}^{-1}(x-a) + \|g(x)\|^2\} f_{\hat{a}}(x) dx
\end{aligned}$$

The first term on the right-hand side of the last equation is independent of $g(x)$. If the second term on the right-hand side is called K, the following can be derived:

$$\begin{aligned}
K &= \int \{2g(x)^T Q_{\hat{a}}^{-1}(x-a) + \|g(x)\|^2\} f_{\hat{a}}(x)dx \\
&= \sum_{z\in\mathbb{Z}^n} \int_{S_z} \{2g(x)^T Q_{\hat{a}}^{-1}(x-a) + \|g(x)\|^2\} f_{\hat{a}}(x)dx \\
&= \sum_{z\in\mathbb{Z}^n} \int_{S_0} \{2g(y)^T Q_{\hat{a}}^{-1}(y+z-a) f_{\hat{a}}(y+z) + \|g(y)\|^2 f_{\hat{a}}(y+z)\} dy \\
&= \int_{S_0} \{2g(y)^T Q_{\hat{a}}^{-1} \underbrace{\sum_{z\in\mathbb{Z}^n} (y+z-a) f_{\hat{a}}(y+z)}_{F(y)} + \|g(y)\|^2 \underbrace{\sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(y+z)}_{G(y)>0}\} dy \\
&= \int_{S_0} \{\underbrace{\|g(y) + \frac{F(y)}{G(y)}\|^2}_{\geq 0} - \underbrace{\|\frac{F(y)}{G(y)}\|^2}_{\text{independent of } g(y)}\} dy
\end{aligned}$$

where S_z are arbitrary pull-in regions, and use is made of the fact that if $x \in S_z$ then $y = x - z \in S_0$ and $g(y+z) = g(y)$.

This shows that K is minimized for $g(y) = -\frac{F(y)}{G(y)}$, which is an admissible solution since $\frac{F(y)}{G(y)}$ is periodic.

□

If $f_{\hat{a}}(x+a)$ is symmetric with respect to the origin, then $\hat{S}(-x) = -\hat{S}(x)$ from which it follows that $\hat{S}(\hat{a})$ is unbiased. In that case the estimator $\hat{S}(\hat{a})$ is of minimum variance and unbiased.

In practice, it is assumed that the float ambiguities are normally distributed. The BIE ambiguity estimator, $\tilde{a}$ follows then from equation (C.3) as:

$$\tilde{a} = \frac{\sum_{z\in\mathbb{Z}^n} z \exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}}{\sum_{z\in\mathbb{Z}^n} \exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}} \tag{C.4}$$

This shows that the BIE ambiguity estimator can be written in a similar form as equation (3.6):

$$\tilde{a} = \sum_{z\in\mathbb{Z}^n} z w_z(\hat{a}) \tag{C.5}$$

The weights $w_z(\hat{a})$ are given by equation (C.4), from which it follows that:

$$0 \leq w_z(\hat{a}) \leq 1 \quad \text{and} \sum_{z\in\mathbb{Z}^n} w_z(\hat{a}) = 1$$

C.4 Best integer equivariant unbiased baseline estimation

In practice the interest is not so much in the ambiguity solution, but in the baseline solution. Let $T(\hat{a}, \hat{b})$ with $F : \mathbb{R}^n \times \mathbb{R}^p \mapsto \mathbb{R}^p$ be the baseline estimator. Two types of restrictions will be put on the function $T(x, y)$. The first requirement is that the if an arbitrary constant $\varsigma \in \mathbb{R}^p$ is added to the float baseline $\hat{b}$, the fixed baseline is shifted by the same amount, i.e.

$$T(x, y + \varsigma) = T(x, y) + \varsigma \quad \forall y, \varsigma \in \mathbb{R}^p \tag{C.6}$$

It follows that if the constant is chosen $\varsigma = -y$ that

$$T(x, y) = y + f(x, 0) \tag{C.7}$$

The second requirement is that the fixed baseline should not change if an arbitrary integer $z \in \mathbb{Z}^n$ is added to the float ambiguities:

$$T(x + z, y) = T(x, y) \quad \forall x \in \mathbb{R}^n, z \in \mathbb{Z}^n \tag{C.8}$$

This is equivalent to the integer remove-restore condition of (C.1).

From (C.7) and (C.8) the following representation of the baseline estimator is found:

$$T(x, y) = y + h(x) \quad \text{with } h(x) \text{ periodic} \tag{C.9}$$

Along similar lines as in section C.3 the BIE baseline estimator can be obtained.

Theorem C.4.1

$$\hat{T}(x, y) = y + b - \frac{\sum\limits_{z \in \mathbb{Z}^n} E\{\hat{b}|x + z\} f_{\hat{a}}(x + z)}{\sum\limits_{z \in \mathbb{Z}^n} f_{\hat{a}}(x + z)} \tag{C.10}$$

is the unique minimizer of $\int\int \|T(x, y) - b\|^2 f_{\hat{a}\hat{b}}(x, y) dx dy$ within the class of integer equivariant estimators, $T(x, y) = y + h(x)$ and $h(x + z) = h(x)$

Proof:

$$\begin{aligned}
&\int\int \|T(x, y) - b\|^2 f_{\hat{a}\hat{b}}(x, y) dx dy \\
&\quad = \int\int \|y + h(x) - b\|^2 f_{\hat{a}\hat{b}}(x, y) dx dy \\
&\quad = \int\int \{\|y - b\|^2 + 2h(x)^T Q^{-1}(y - b) + \|h(x)\|^2\} f_{\hat{a}\hat{b}}(x, y) dx dy
\end{aligned}$$

The first term on between the braces on the right-hand side of the last equation is independent of $h(x)$. If the remaining part is called L and $q = x - z \in S_0$, the following

can be derived:

$$
\begin{aligned}
L &= \int \sum_{z\in\mathbb{Z}^n} \int \{2h(x)^T Q^{-1}(y-b) + \|h(x)\|^2\} f_{\hat{b}|\hat{a}}(y|x) f_{\hat{a}}(x) dx dy \\
&= \int \int_{S_0} \{2h(q)^T Q^{-1}(y-b) + \|h(q)\|^2\} \sum_{z\in\mathbb{Z}^n} f_{\hat{b}|\hat{a}}(y|q+z) f_{\hat{a}}(q+z) dq dy \\
&= \int_{S_0} \{2h(q)^T Q^{-1} \sum_{z\in\mathbb{Z}^n} \int (y-b) f_{\hat{b}|\hat{a}}(y|q+z) dy \\
&\qquad\qquad + \|h(q)\|^2\} \sum_{z\in\mathbb{Z}^n} \int f_{\hat{b}|\hat{a}}(y|q+z) dy\} f_{\hat{a}}(q+z) dq \\
&= \int_{S_0} \{2h(q)^T Q^{-1} \underbrace{\sum_{z\in\mathbb{Z}^n} (E\{\hat{b}|q+z\} - b) f_{\hat{a}}(q+z)}_{F(q)} + \|h(q)\|^2 \underbrace{\sum_{z\in\mathbb{Z}^n} f_{\hat{a}}(q+z)}_{G(q)}\} dq \\
&= \int_{S_0} \{\|h(q) + \frac{F(q)}{G(q)} - b\|^2 - \|\frac{F(q)}{G(q)} - b\|^2\} G(q) dq
\end{aligned}
$$

L is minimized for $h(q) = -\frac{F(q)}{G(q)} + b$, which is an admissible solution since $\frac{F(q)}{G(q)}$ is periodic.

□

Note that $\hat{T}$ in (C.10) is the BIE unbiased estimator if $f_{\hat{a}}(x+a)$ is symmetric with respect to the origin.

If the float estimators are normally distributed the BIE baseline estimator, $\tilde{b}$, follows from equation (C.10) as:

$$
\begin{aligned}
\tilde{b} &= \hat{b} + b - \frac{\sum_{z\in\mathbb{Z}^n} E\{\hat{b}|\hat{a}+z\} \exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}}{\sum_{z\in\mathbb{Z}^n} \exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}} \\
&= \hat{b} + b - \frac{\sum_{z\in\mathbb{Z}^n} \left(b + Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(\hat{a}-z)\right) \exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}}{\sum_{z\in\mathbb{Z}^n} \exp\{-\frac{1}{2}\|x-z\|^2_{Q_{\hat{a}}}\}} \\
&= \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(\hat{a}-\tilde{a})
\end{aligned}
\tag{C.11}
$$

where the BIE ambiguity estimator of equation (C.4) is used.

Implementation aspects of IALS estimation

D

The procedure outlined in section 5.7.3 can also be applied for the IALS estimator, but is more complicated as will be explained here. The problem is the determination of the aperture parameter such that the float ambiguity estimate will be on the boundary of the aperture pull-in region. From the definition of the ILS pull-in region in equation (3.17) follows that $\hat{x}$ will be on the boundary of $S_{\check{x}}$ if:

$$\left| \frac{(\check{x}_2 - \check{x})^T Q_{\hat{x}}^{-1} (\hat{x} - \check{x})}{\|\check{x}_2 - \check{x}\|_{Q_{\hat{x}}}^2} \right| = \frac{1}{2} \tag{D.1}$$

with $\check{x}$ and $\check{x}_2$ the best and second-best integers in the integer least-squares sense.

From the definition of the IALS estimator in section 5.4.2 it follows that $\hat{x} - \check{x}$ is on the boundary of the aperture pull-in region Ω_0 if $\frac{1}{\mu}(\hat{x} - \check{x})$ is on the boundary of S_0. Furthermore, the following will be used:

$$\begin{cases} \arg \min\limits_{z \in \mathbb{Z}^n} \|\frac{1}{\mu}(\hat{x} - \check{x}) - z\|_{Q_{\hat{x}}}^2 = 0 \\ \arg \min\limits_{z \in \mathbb{Z}^n \setminus \{0\}} \|\frac{1}{\mu}(\hat{x} - \check{x}) - z\|_{Q_{\hat{x}}}^2 = u \end{cases} \tag{D.2}$$

These results are used in equation (D.1) to find an expression for μ:

$$\left| \frac{u^T Q_{\hat{x}}^{-1} \frac{1}{\mu}(\hat{x} - \check{x})}{\|u\|_{Q_{\hat{x}}}^2} \right| = \frac{1}{2} \quad \Longrightarrow \quad \mu = \left| \frac{2u^T Q_{\hat{x}}^{-1} (\hat{x} - \check{x})}{\|u\|_{Q_{\hat{x}}}^2} \right| \tag{D.3}$$

The problem is that u can only be determined if μ is known. Therefore as a first guess, one could use $u^0 = \check{x}_2 - \check{x}$ and then compute the corresponding μ^0 with equation (D.3) by using u^0 instead of u. If

$$\arg \min_{z \in \mathbb{Z}^n \setminus \{0\}} \|\frac{1}{\mu^0}(\hat{x} - \check{x}) - z\|_{Q_{\hat{x}}}^2 = u \tag{D.4}$$

and $u = \check{x}_2 - \check{x}$ it means that indeed $\mu^0 = \mu$. If, on the other hand, $u \neq \check{x}_2 - \check{x}$ it means that although $\frac{1}{\mu^0}(\hat{x} - \check{x})$ is on the plane that is bounding S_0, this vector is not on the boundary of S_0. Hence, $\frac{1}{\mu^0}(\hat{x} - \check{x}) \in S_u$. This implies that $\mu^0 < \mu$ and that $\frac{1}{\mu}(\hat{x} - \check{x})$ is on the boundary of S_u and S_0. Now that u is known, μ can finally be computed with equation (D.3).

The procedure for IALS estimation is outlined below.

1. Choose fixed fail rate: $P_f = \beta$
2. Collect observations and apply least-squares adjustment: $\hat{a}$, $Q_{\hat{a}}$;
3. Apply integer least-squares estimation: $\check{a}$, $\check{\epsilon}$, $\check{a}_2$;
4. Choose μ' as described above;
5. Generate N samples of float ambiguities $\hat{x}_i \sim N(0, Q_{\hat{a}})$;
6. For each sample determine the ILS estimate $\check{x}_i$ and $u_i = \arg\min_{z \in \mathbb{Z}^n} \| \frac{1}{\mu}(\hat{x}_i - \check{x}_i) - z \|^2_{Q_{\hat{x}}}$;
7. Count number of samples N_f for which: $u_i = 0$ and $\check{x}_i \neq 0$;
8. Compute fail rate with μ' as critical value: $P_f(\mu') = \frac{N_f}{N}$;
9. If $P_f(\mu') \leq \beta$: $\bar{a} = \check{a}$, otherwise $\bar{a} = \hat{a}$;

Note that in step 6 ILS estimation has to be carried out twice for each sample. This might be very time-consuming if the number of samples, N, is large.

E

Curriculum vitae

Sandra Verhagen (1976) was born in Leidschendam, The Netherlands.

1988 - 1994	Alfrink College, Zoetermeer
1994 - 2000	Geodetic Engineering, Delft University of Technology Thesis: *Time variations in the gravity field - The effect of the atmosphere*
2000 - 2004	PhD researcher, Mathematical Geodesy and Positioning
2003	ESA project *Analysis of ambiguity resolution algorithms and quality control parameters for Global Navigation Satellite Systems*

Bibliography

Abidin HA (1993). *Computational and geometrical aspects of on-the-fly ambiguity resolution*. Ph.D. Thesis, Dept. of Surveying Engineering, Techn.Report no.104, University of New Brunswick, Canada, 314 pp.

Baarda W (1968). *A testing procedure for use in geodetic networks*. Publications on Geodesy, 2(5), Netherlands Geodetic Commission, Delft.

Baarda W (1973). *S-transformations and criterion matrices*. Publications on Geodesy, 5(1), Netherlands Geodetic Commission, Delft.

Betti B, Crespi M, Sansò F (1993). A geometric illustration of ambiguity resolution in GPS theory and a bayesian approach. *Manuscripta Geodaetica*, 18: 317–330.

Blewitt G (1989). Carrier-phase ambiguity reolution for the Global Positioning System applied to baselines up to 2000 km. *Journal of Geophysical Research*, 94(B8): 10187–10302.

Bona P (2000). Precision, cross corelation, and time correlation of GPS phase and code observations. *GPS Solutions*, 4(2): 3–13.

Bona P, Tiberius CCJM (2001). An experimental assessment of observation cross-correlation in dual frequency GPS receivers. *Proc. of ION GPS-2000, Salt Lake City UT*: 792–798.

Borre K, Tiberius CCJM (2000). Time series analysis of GPS observables. *Proc. of ION GPS-2000, Salt Lake City UT*: 1885–1894.

Brent R (1973). *Algorithms for minimization without derivatives*. Prentice-Hall.

Chen D, Lachapelle G (1995). A comparison of the FASF and least-squares search algorithms for on-the-fly ambiguity resolution. *Navigation*, 42(2): 371–390.

Chen Y (1997). An approach to validate the resolved ambiguities in GPS rapid positioning. *Proc. of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation, Banff, Canada*.

De Jonge PJ, Tiberius CCJM (1996). *The LAMBDA method for integer ambiguity estimation: implementation aspects*. Delft Geodetic Computing Centre, LGR series No.12, Delft University of Technology, 49pp.

Dong D, Bock Y (1989). Global Positioning System network analysis with phase ambiguity resolution applied to crustal deformation studies in California. *Journal of Geophysical Research*, 94(B4): 3949–3966.

Erhard P (2002). *Status of Galileo navigation signals, Design and frequency plan*. ESA publication "ESA-APPNG-TN/00407/PhE", 59 p.

Euler HJ, Goad C (1991). On optimal filtering of GPS dual frequency observations without using orbit information. *Bulletin Géodésique*, 65: 130–143.

Euler HJ, Landau H (1992). Fast GPS ambiguity resolution on-the-fly for real-time application. *Proc. of sixth International Geodetic Symposium on Satellite Positioning, Columbus OH*: 650–659.

Euler HJ, Schaffrin B (1991). On a measure for the discernibility between different ambiguity solutions in the static-kinematic GPS-mode. *IAG Symposia no.107, Kinematic Systems in Geodesy, Surveying, and Remote Sensing, Springer-Verlag, New York*: 285–295.

Fernàndez-Plazaola U, Martín-Guerrero TM, Entrambasaguas-Muñoz JT, Martín-Neira M (2004). The Null method applied to GNSS three-carrier phase ambiguity resolution. *Journal of Geodesy*, 78: 96–102.

Frei E, Beutler G (1990). Rapid static positioning based on the fast ambiguity resolution approach FARA: theory and first results. *Manuscripta Geodaetica*, 15: 325–356.

Gundlich B (2002). *Statistische Untersuchung ganzzahliger und reellwertiger unbekannter Parameter im GPS-Modell*. Ph.D. thesis, DGK, Reihe C, no. 549, Muenchen.

Gundlich B, Koch KR (2002). Confidence regions for GPS baselines by Bayesian statistics. *Journal of Geodesy*, 76: 55–62.

Han S (1997). Quality control issues relating to instantaneous ambiguity resolution for real-time GPS kinematic positioning. *Journal of Geodesy*, 71: 351–361.

Han S, Rizos C (1996a). Integrated methods for instantaneous ambiguity resolution using new-generation GPS receivers. *Proc. of IEEE PLANS'96, Atlanta GA*: 254–261.

Han S, Rizos C (1996b). Validation and rejection criteria for integer least-squares estimation. *Survey Review*, 33(260): 375–382.

Harris RA (1997). *Direct resolution of carrier-phase ambiguity by 'Bridging the wavelength gap'*. ESA publication "TST/60107/RAH/Word".

Hassibi A, Boyd S (1998). Integer parameter estimation in linear models with applications to GPS. *IEEE Transactions on Signal Processing*, 46(11): 2938–2952.

Hatch R (1990). Instantaneous ambiguity resolution. *Proc. of KIS'90, Banff, Canada*: 290–308.

Hofmann-Wellenhof B, Lichtenegger H, Collins J (2001). *Global Positioning System: Theory and Practice* (5th ed.). Springer-Verlag, Berlin.

Hopfield HS (1969). Two-quartic tropospheric refractivity profile for correcting satellite data. *Journal of Geophysical Research*, 74(18): 4487–4499.

Jonkman NF (1998). *Integer GPS ambiguity estimation without the receiver-satellite geometry*. Delft Geodetic Computing Centre, LGR series No.18, Delft University of Technology, 95pp.

Joosten P, Pany T, Winkel J (2002). The impact of unmodelled multipath on ambiguity resolution. *Proc. of ION GPS-2002, Portland OR*: 953–959.

Joosten P, Verhagen S (2003). *Analysis of ambiguity resolution algorithms and quality control parameters for Global Navigations Satellite Systems*. ESA publication, contract no. "16793/02/NL/LvH".

Jung J, Enge P, Pervan B (2000). Optimization of Cascade Integer Resolution with three civil GPS frequencies. *Proc. of ION GPS-2000, Salt Lake City UT*: 2191–2200.

Kim D, Langley RB (1999). An optimized least-squares technique for improving ambiguity resolution performance and computational efficiency. *Proc. of ION GPS-1999, Nashville TN*: 1579–1588.

Kim D, Langley RB (2000). GPS ambiguity resolution and validation: methodologies, trends and issues. *Proc. of 7th GNSS Workshop, Seoul, Korea*.

Kleijer F (2004). *Troposphere modeling and filtering for precise GPS leveling*. Ph.D. thesis, Publications on Geodesy, 56, Netherlands Geodetic Commission, Delft.

Kondo K (2003). Optimal success/error rate and its calculation in resolution of integer ambiguities in carrier phase positioning of Global Positioning System (GPS) and Global Navigation Satellite System (GNSS). *Proc. of ION Annual Meeting, Albuquerque NM*: 176–187.

Landau H, Euler HJ (1992). On-the-fly ambiguity resolution for precise differential positioning. *Proc. of ION GPS-1992, Albuquerque NM*: 607–613.

Leick A (2003). *GPS Satellite Surveying* (3rd ed.). John Wiley and Sons, New York.

Liu X (2002). A comparison of stochastic models for GPS single differential kinematic positioning. *Proc. of ION GPS-2002, Portland OR*: 1830–1841.

Martín-Neira M, Toledo M, Pelaez A (1995). The null space method for GPS integer ambiguity resolution. *Proc. of DSNS'95, Bergen, Norway, paper no.31*.

Neilan RE, Moore A, Springer T, Kouba J, Ray J, Reigber C (2000). International GPS Service 2000: Life without SA. *Proc. of ION GPS-2000, Salt Lake City UT*: 438–446.

Odijk D (2002). *Fast precise GPS positioning in the presence of ionospheric delays*. Ph.D. thesis, Publications on Geodesy, 52, Netherlands Geodetic Commission, Delft.

Parkinson BW, Spilker JJ (Eds.) (1996). *Global Positioning System: Theory and Applications, Vols. 1 and 2*. Volume 164 of Progress in Aeronautics and Astronautics, AIAA, Washington DC.

Rao CR (1965). *Statistical inference and its applications*. John Wiley and Sons, New York.

Saastamoinen J (1973). Contributions to the theory of atmospheric refraction. *Bulletin Géodésique*, 105: 279-298, 106: 383-397, 107: 13-34.

Salgado G, Abbondanza S, Blondel R, Lannelongue S (2001). Constellation availability concepts for Galileo. *Proc. of ION NTM-2001, Long Beach CA*: 778–786.

Strang G, Borre K (1997). *Linear Algebra, Geodesy, and GPS*. Wellesley-Cambridge Press, Wellesley MA.

Teunissen PJG (1984). Generalized inverses, adjustment, the datum problem and S-transformations. *Lecture notes International School of Geodesy, Erice, Italy, April 25 - May 10*.

Teunissen PJG (1990). An integrity and quality control procedure for use in multi-sensor integration. *Proc. of ION GPS-1990, Colorado Springs CO*: 513–522.

Teunissen PJG (1993). Least squares estimation of the integer GPS ambiguities. *Invited lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing*.

Teunissen PJG (1995). The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, 70: 65–82.

Teunissen PJG (1997). A canonical theory for short GPS baselines. Part IV: Precision versus reliability. *Journal of Geodesy*, 71: 513–525.

Teunissen PJG (1998a). A class of unbiased integer GPS ambiguity estimators. *Artificial Satellites*, 33(1): 4–10.

Teunissen PJG (1998b). *GPS carrier phase ambiguity fixing concepts*. In: PJG Teunissen and A Kleusberg, *GPS for Geodesy*, Springer-Verlag, Berlin.

Teunissen PJG (1998c). On the integer normal distribution of the GPS ambiguities. *Artificial Satellites*, 33(2): 49–64.

Teunissen PJG (1998d). Success probability of integer GPS ambiguity rounding and bootstrapping. *Journal of Geodesy*, 72: 606–612.

Teunissen PJG (1998e). The distribution of the GPS baseline in case of integer least-squares ambiguity estimation. *Artificial Satellites*, 33(2): 65–75.

Teunissen PJG (1999a). An optimality property of the integer least-squares estimator. *Journal of Geodesy*, 73: 587–593.

Teunissen PJG (1999b). The probability distribution of the GPS baseline for a class of integer ambiguity estimators. *Journal of Geodesy*, 73: 275–284.

Teunissen PJG (2000a). *Adjustment theory; an introduction*. Delft University Press, Delft.

Teunissen PJG (2000b). ADOP based upperbounds for the bootstrapped and the least-squares ambiguity success rates. *Artificial Satellites*, 35(4): 171–179.

Teunissen PJG (2001a). Integer estimation in the presence of biases. *Journal of Geodesy*, 75: 399–407.

Teunissen PJG (2001b). Statistical GNSS carrier phase ambiguity resolution: a review. *Proc. of 2001 IEEE Workshop on Statistical Signal Processing, August 6-8, Singapore*: 4–12.

Teunissen PJG (2002). The parameter distributions of the integer GPS model. *Journal of Geodesy*, 76: 41–48.

Teunissen PJG (2003a). A carrier phase ambiguity estimator with easy-to-evaluate fail-rate. *Artificial Satellites*, 38(3): 89–96.

Teunissen PJG (2003b). An invariant upperbound for the GNSS bootstrapped ambiguity success rate. *Journal of Global Positioning Systems*, 2(1): 13–17.

Teunissen PJG (2003c). GNSS Best Integer Equivariant Estimation. *Accepted for publication in: Proc. of IUGG2003, session G.04, June 30 - July 11, Sapporo, Japan*.

Teunissen PJG (2003d). Integer aperture GNSS ambiguity resolution. *Artificial Satellites*, 38(3): 79–88.

Teunissen PJG (2003e). *Theory of integer aperture estimation with application to GNSS*. MGP report, Delft University of Technology.

Teunissen PJG (2003f). Theory of integer equivariant estimation with application to GNSS. *Journal of Geodesy*, 77: 402–410.

Teunissen PJG (2003g). Towards a unified theory of GNSS ambiguity resolution. *Journal of Global Positioning Systems*, 2(1): 1–12.

Teunissen PJG (2004a). Integer aperture bootstrapping: A new GNSS ambiguity estimator with controllable fail-rate. *Submitted to Journal of Geodesy*.

Teunissen PJG (2004b). Integer Aperture Least-Squares estimation. *Submitted to Artificial Satellites*.

Teunissen PJG (2004c). On the computation of the BIE estimator. *Submitted to Artificial Satellites*.

Teunissen PJG (2004d). Optimal Integer Aperture estimation. *Submitted to Artificial Satellites*.

Teunissen PJG (2004e). Penalized GNSS ambiguity resolution. *Journal of Geodesy, DOI: 10.1007/s00190-004-0393-2*.

Teunissen PJG, De Jonge PJ, Tiberius CCJM (1996). The volume of the GPS ambiguity ambiguity search space and its relevance for integer ambiguity resolution. *Proc. of ION GPS-1996, Kansas City MO*: 889–898.

Teunissen PJG, Joosten P, Tiberius CCJM (2000). Bias robustness of GPS ambiguity resolution. *Proc. of ION GPS-2000, Salt Lake City UT*: 104–112.

Teunissen PJG, Kleusberg A (1998). *GPS observation equations and positioning concepts*. In: PJG Teunissen and A Kleusberg, *GPS for Geodesy*, Springer-Verlag, Berlin.

Teunissen PJG, Odijk D (1997). Ambiguity Dilution of Precision: definition, properties and application. *Proc. of ION GPS-1997, Kansas City MO*: 891–899.

Tiberius CCJM, De Jonge PJ (1995). Fast positioning using the LAMBDA method. *Proc. of DSNS'95, Bergen, Norway, paper no.30*.

Tiberius CCJM, Kenselaar F (2000). Stochastic model for GPS code and phase observables. *Survey Review*, 35: 441–454.

Verhagen S (2003). On the approximation of the integer least-squares success rate: which lower or upper bound to use? *Journal of Global Positioning Systems*, 2(2): 117–124.

Verhagen S (2004). Integer ambiguity validation: an open problem? *GPS Solutions*, 8(1): 36–43.

Verhagen S, Joosten P (2004). Analysis of ambiguity resolution algorithms. *Proc. of the European Navigation Conference GNSS 2004, Rotterdam NL*.

Verhagen S, Teunissen PJG (2003). Performance comparison of the BIE estimator with the float and fixed GNSS ambiguity estimators. *Accepted for publication in: Proc. of IUGG2003, session G.04, June 30 - July 11, Sapporo, Japan*.

Verhagen S, Teunissen PJG (2004a). A new GNSS ambiguity resolution method and its comparison with existing approaches. *Submitted to Journal of Guidance, Control, and Dynamics*.

Verhagen S, Teunissen PJG (2004b). On the probability density function of the GNSS ambiguity residuals. *Submitted to GPS Solutions*.

Verhagen S, Teunissen PJG (2004c). *PDF evaluation of the ambiguity residuals*. In: F Sansò (Ed.), *V. Hotine-Marussi Symposium on Mathematical Geodesy*, International Association of Geodesy Symposia, Vol. 127, Springer-Verlag.

Vollath U, Birnbach S, Landau H, Fraile-Ordoñez JM, Martín-Neira M (1998). Analysis of Three-Carrier Ambiguity Resolution (TCAR) Technique for precise relative positioning in GNSS-2. *Proc. of ION GPS-1998, Nashville TN*: 417–426.

Wang J, Stewart MP, Tsakiri M (1998). A discrimination test procedure for ambiguity resolution on-the-fly. *Journal of Geodesy*, 72: 644–653.

Wei M, Schwarz KP (1995). Fast ambiguity resolution using an iteger nonlinear programming method. *Proc. of ION GPS-1995, Palm Springs CA*: 1101–1110.

Zandbergen R, Dinwiddy S, Hahn J, Breeuwer E, Blonski D (2004). Galileo orbit selection. *Proc. of ION GPS-2004, September 21-24, Long Beach CA*.

Index